UNDERSTANDING SIGNALS WITHIN THE ONLINE STOLEN DATA MARKET: AN EXAMINATION OF VENDORS' SIGNALING BEHAVIORS RELATIVE TO STOLEN DATA PRICE POINTS

By

Jin Ree Lee

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Criminal Justice — Doctor of Philosophy

2021

ABSTRACT

UNDERSTANDING SIGNALS WITHIN THE ONLINE STOLEN DATA MARKET: AN EXAMINATION OF VENDORS' SIGNALING BEHAVIORS RELATIVE TO STOLEN DATA PRICE POINTS

By

Jin Ree Lee

Over the last 15 years, research has demonstrated that a robust, global illicit marketplace exists for the sale of personal information acquired through data breaches and other forms of hacking. There is particular emphasis on the sale of financial data (e.g., credit/debit card numbers) as this information can be used to commit various forms of economic crime, including credit card fraud and identity theft. Identity-related offenses involving stolen financial data are both the most common and fastest growing forms of consumer fraud on the Internet and engender significant economic harm to corporations and individual consumers alike. Despite the constant state of data breaches and growing number of illicit stolen data markets, few have examined the pricing structure of online financial data products to assess vendor unreliability. Research suggests stolen data vendors use the price point for a given item as a signaling mechanism to express their credibility. The lack of price structure analyses on stolen data products may be a result of the complexities involved in accurately measuring products' price point. In addition, few have explicitly examined stolen data vendors' signaling practices in relation to products' price point on both the Open and Dark Web, as well as across different types of markets within the same analysis. Understanding the relationship between signaling behaviors and vendor unreliability is important as it could help investigative operations differentiate substantive mechanisms and practices from less concerning noise in terms of identifying serious market actors from those that pose lower levels of threat. Relatedly, a deeper understanding of vendors' signaling behaviors at both the product and vendor-level would assist law enforcement devise effective intervention strategies targeted at disrupting the online illicit marketplace. From a theoretical perspective, such research also extends understanding of traditional criminological theory by illustrating its ability to explain emerging forms of crime and deviance. Given this gap in the literature, the current study explored the signaling mechanisms of 1,055 stolen data products across 40 vendors on the Open (n = 8; 20%) and Dark Web (n = 32; 80%). Specifically, the current study used a signaling theory framework to examine whether vendors' differential use of product-level (e.g., detailed product information) and vendor-level indicators (e.g., payment methods, customer service features, customer feedback mechanisms) predicted stolen data price points. The research and policy implications of this study in understanding the signaling behaviors of online stolen data vendors are explored in detail.

ACKNOWLEDGEMENTS

I would like to preface this section by giving glory to my Lord and Savior, Jesus Christ. My journey of becoming a Ph.D. would not have been possible without the guidance, support, and prayer of an entire community. First, I would like to express my deepest gratitude to my dissertation committee: Drs. Thomas J. Holt, Steven M. Chermak, Scott E. Wolfe, and George W. Burruss. Thank you for your insightful comments and thoughtful feedback throughout this entire process. I am forever indebted to your unwavering support, mentorship, and guidance. I would also like to extend my gratitude to Drs. Karen M. Holt, Christina DeJong, Jennifer E. Cobbina, Hyeyoung Lim, Sameer Hinduja, and Justin W. Patchin for always encouraging me to be the best I can be. Your undivided attention and ceaseless support gave me the strength I needed to persevere and overcome life's many obstacles and challenges.

I would also like to thank my past advisors and mentors for encouraging me to pursue my dreams. Thank you, Dr. Steven Downing, for taking a chance on me and encouraging me to chase after my goals. I am forever indebted to your mentorship. Thank you, Dr. Barbara Perry, for teaching me how to navigate the academy and challenging me to become a better writer. Thank you, Dr. Andrea Slane, for exemplifying true mentorship and teaching me how to become a devoted researcher.

Finally, I would like to express my sincerest gratitude to a handful of individuals who have been pivotal in my life. First and foremost: Thank you, Dr. Thomas J. Holt, for being my Ph.D. mentor and serving as my dissertation advisor. Thank you for believing in me before I believed in myself. If I could end up being half the scholar and mentor that you are, I would

consider my life well-lived. I am forever honored to be your student, mentee, and friend. Thank you for everything, always.

To the greatest collection of colleagues ever—Jacek Koziarski, David Y. Nam, Ariel L. Roddy, Brenna M. Helm, Roberta E. Liggett-O'Malley, Kathleen M. Darcy, Rebecca Malinski, Brent R. Klein, Christine Kwiatkowski, Noah D. Turner, Lisa Moore, Alison Cox, Spencer G. Lawson, Jennifer Paruk, Rachel Boratto, Fiona Chan, Mikaela Wallin, Kayla Hoskins, Kourtnie Rodgers, and Michelle Malkin—thank you for enduring this long and tireless journey by my side. I am eternally grateful for you all. Thank you for being my constant source of joy, energy, and inspiration.

Lastly, I would like to thank my dearest family for their boundless love and support. To my beloved parents, Rev. Jay J.K. Lee and Sunny Y.S. Lee: Thank you for your endless sacrifices in raising me to become something more than I had ever imagined. Thank you for putting me and my aspirations above yourselves and for your ceaseless prayers. You gave up so much of yourselves so that I could have the world. I am forever grateful for your love, kindness, and support. To my sister, Bo Bae Lee: Thank you for always praying for me and encouraging me to be better than my yesterday. You are a true inspiration to this world. To my parents-in-law, Dr. Shin-Ho Lee and Mi-Ae Kwon: Thank you for always blessing me with love, prayer, and wisdom. I am eternally grateful for your love and support. Thank you for always watching over me with prayer and encouraging me to be the best I can be. Last, but certainly not least, to my wonderful bride: Yoon Jeong, thank you for your undivided attention, unrelenting patience, and endless grace. Thank you for being the greatest presence in my life. None of this would have been possible without your love and support. I am forever grateful for your companionship, love, and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The growth and ubiquity of the Internet has substantially modified the way personal and financial data are accessed and maintained, providing individuals with fast and easy access to electronic funds and financial accounts (James, 2005; Newman & Clarke, 2003; Wall, 2001). E-commerce platforms and financial portals allow individuals to engage in instant economic transactions as long as they have access to an Internet-enabled device (e.g., computer, smartphone, tablet computer) and high-speed Internet, influencing the way individuals, businesses, and governments interact with one another (Holt, 2020; James, 2005; Newman & Clarke, 2003; Wall, 2001). According to the United States Department of Commerce, online retail has emerged as the fourth largest sector in retail spending, trailing only motor vehicle, food and beverage, and restaurant and bar sales (Rooney, 2019). In fact, the online retail industry has amounted to approximately $3.53 trillion in global sales (Clement, 2019), with 80% of Americans reporting making regular purchases online (Smith & Anderson, 2016).

Despite its overall convenience and accessibility as a platform, the growth of e-commerce and concurrent decentralization of private financial data (e.g., credit and bank card information) has generated numerous opportunities for offenders to access, steal, and abuse sensitive financial information (Franklin, Perrig, Paxson, & Savage, 2007; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010; Holt, Smirnova, & Chua, 2016; James, 2005; Newman & Clarke, 2003; Peretti, 2009). Numerous major retailers (e.g., Home Depot, Target, Under Armour), travel and hospitality agencies (e.g., EasyJet, Marriott, Norwegian Cruise Line), and restaurant chains (e.g., Bubba Gump Shrimp, Dunkin' Donuts, P.F. Chang's) had their financial databases compromised by hackers in recent years, resulting in the loss of millions of customers' data (National Technology Security Coalition, 2020; Ponemon Intitute, 2019; SelfKey, 2021; Verizon, 2020).

The number of global data breaches affecting critical infrastructure, public corporations, and government agencies has risen exponentially over the past two decades, generating significant economic loss to both organizations and individuals (Higgins, 2014; Huang & Ban, 2019; National Technology Security Coalition, 2020; Ponemon Institute, 2019; Seals, 2014; SelfKey, 2021; Verizon, 2020). A data breach is generally defined as an incident where protected sensitive information (e.g., social security numbers, bank account or credit card numbers, email and account passwords, and personal health information) is unlawfully accessed by an unauthorized actor(s) (see Higgins, 2014; Holt & Bossler, 2016; Ponemon Institute, 2019; Seals, 2014; Verizon, 2020). In fact, IBM (2020) estimated the average cost of a global data beach to be around $3.9 million (see also Franklin et al., 2007; Holt, Smirnova, & Chua, 2016; Holz, Engelberth, & Freiling, 2009; Moore, Clayton, & Anderson, 2009). The immediate financial ramifications of data breaches are concerning for corporate victims, such as legal costs, share slumps, and remediation/repair costs. The long-term consequences may be even more catastrophic due to losses in customer confidence and broader harm to corporate competitiveness (Wilson, 2019).

The individual victims whose financial data were compromised are, however, the ones most negatively affected by corporate data breaches. Individual consumers are more likely to experience financial losses, as well as direct secondary impacts, including experiencing stress, anxiety, and having to devote time and resources resolving the problems caused by the data breach (Holt & Bossler, 2016; Holt, Smirnova, & Chua, 2016; Ponemon Institute, 2014). In fact, the Ponemon Institute (2014) found 76% of individual victims of corporate data breaches experienced stress as a result of their data being unlawfully accessed or compromised, while

39% had to spend time repairing the damages caused by the data breach (Ponemon Institute, 2014).

There are many ways for economically motivated offenders to acquire large sums of sensitive financial data from online targets, including computer virus infections and social engineering schemes (e.g., phishing and spam emails) (see Furnell, 2002; Holt, 2020; Holt & Bossler, 2016; Huang & Brockman, 2010; Hutchings & Holt, 2017; James, 2005; Mitnick & Simon, 2002; Ngo, Agarwal, Govindu, & MacDonald, 2020; Peretti, 2009). The quantity of financial data that can be acquired by these actors at a single time often exceeds their capacity for personal use (Holt, Smirnova, & Hutchings, 2016). As a result, surplus financial data are commonly sold and distributed via illicit markets operating on various forms of computer-mediated communications (e.g., online stolen data markets) to other individuals interested in using the data for their own advantage (Franklin et al., 2007; Holt, Smirnova, Chua, & Copes, 2015; Holt, Smirnova, & Hutchings, 2016; Holz, Engelberth, & Freiling, 2009; Motoyama, McCoy, Levchenko, Savage, & Voelker, 2011; Yip, Webber, & Shadbolt, 2013).

Stolen data vendors offer a range of products at low costs, including credit and debit card numbers, card verification values (CVVs), and cardholders' related information (Anderson et al., 2013; Decary-Hetu & Laferriere, 2015; Franklin et al., 2007; Holt & Bossler, 2016; Holt & Lampke, 2010; Liggett et al., 2020). Depending on the product's contained amount, card type, and quantity, items can be advertised as low as $1 per individual product (see Holt, Smirnova, & Chua, 2016). Some vendors even offer customers the option to purchase financial data in either digital form as raw data or as physical cards for use in stores (Decary-Hetu & Leppanen, 2016; Franklin et al., 2007; Holt & Lampke, 2010; Holz, Engelberth, & Freiling, 2009).

The ascent of the stolen data marketplace is unsurprising given historical evidence that hackers monetized the spoils of hacking in different ways, such as trading pirated software, login IDs, and passwords (Dupont, Cote, Boutin, & Fernandez, 2017; Holt & Lampke, 2010; Holt et al., 2015; Hutchings & Holt, 2015; Motoyama et al., 2011). The growth of the stolen data market can also be ascribed in part to shifts in the behaviors of hackers and data thieves over the last two decades (see Holt, 2020). Computer hacking has transitioned from being a behavior primarily dominated by highly skilled individuals, to a market economy where less technologically proficient individuals are able to purchase pre-assembled tools and kits from skilled actors on a pay-per-service basis (Cooke, Jahanian, & McPherson, 2005; Decary-Hetu & Dupont, 2013; Hyslip, 2020; Hyslip & Holt, 2019). Tools can range from pre-made malicious computer software (e.g., worms, trojans, viruses) to fully programmed attack tools (e.g., bots, botnet) that can be operated to compromise a target's system or database. Pre-packaged malicious software and automated attack tools are advertised through various online forums and shops for interested buyers to purchase, often indicating the product's functionality (i.e., its properties, limitations, and capabilities) and instructions on how it can be operated to successfully generate an attack (Holt, 2007; 2017; Holt, Bossler, & Seigfried-Spellar, 2018). These products and services enable non-technical offenders to target systems even if they do not possess the technical understanding and expertise needed to create and execute an attack on their own.

Though online stolen data markets represent cybercrimes that involve the sale of private data acquired through somewhat sophisticated methods, they are committed by individuals who are motivated by many of the same factors observed with traditional offline crimes and are inherently driven by the same human decision-making process (see Leukfeldt & Holt, 2019). Given the rise of online stolen data markets, a better understanding of the situational processes that shape the decisions of financially motivated hackers and data thieves is needed. To that end, classical perspectives in criminology have been used to understand individuals' involvement in various economic behaviors, including drug crimes (Hamid, 1998; Jacobs, 1996a, 2000; Johnson, Dunlap, & Tourigny, 2000; Johnson & Natarajan, 1995; Knowles, 1999; Lupton, Wilson, May, Warburton, & Turnbull, 2002; McSweeney, Turnbull, & Hough, 2008; Topalli, Wright, & Fornango, 2002; VanNostrand & Tewksbury, 1999) and prostitution offenses (Holt, Blevins, & Kuhns, 2014; Scott & Dedel, 2006).

Classical perspectives generally argue that individuals are rational actors motivated by their self-interests to maximize pleasure and minimize pain (Becker, 1968; Clarke, 1997; Clarke & Cornish, 1985; Cullen, Agnew, & Wilcox, 2014; Holt & Dupont, 2019). Rational choice and deterrence theories suggest people engage in criminal conduct if the pleasures associated with the behavior outweigh the pains associated with the act (Becker, 1968; Clarke, 1997; Clarke & Cornish, 1985; Cullen, Agnew, & Wilcox, 2014; Holt & Dupont, 2019). Consequently, criminal behavior is deterred by both increasing the likelihood of detection (e.g., certainty of punishment) and reducing the expected benefit of the act (Cornish & Clarke, 1987; Cullen, Agnew, & Wilcox, 2014).

Rational choice theory is particularly valuable to examine illicit markets and their operational practices, as it assumes actors make calculated decisions as to which markets, vendors, and products are sought out to maximize their overall pleasure and minimize risk (see Becker, 1968; Clarke, 1997; Clarke & Cornish, 1985; Holt & Dupont, 2019). Specifically, illicit vendors conduct cost-benefit analyses where they balance the benefits of operating a profitable illegal business, with the risks of being harmed by competing sellers and/or apprehended by law enforcement (Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt, Smirnova, &

Chua, 2016; Holt, Smirnova, & Hutchings, 2016; Hutchings & Holt, 2017). Similarly, prospective customers measure the benefits of purchasing a desired product from a particular vendor, with the risks of doing business with that specified seller and product (Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Chua, 2016; Holt, Smirnova, & Hutchings, 2016; Hutchings & Holt, 2017).

The scope of data breaches against businesses has generated a highly saturated and competitive stolen data marketplace, offering prospective customers a wide variety of vendor options and product choices (see Holt, Chua, & Smirnova, 2013; Tzanetakis, Kamphausen, Werse, & von Laufenberg, 2016). Individuals can now buy and sell stolen financial data on the Open Web, generally described as the widely accessible part of the web that can be accessed by conventional browsers (e.g., Google Chrome, Firefox, Internet Explorer) and search engines (e.g., Google, Yahoo, Bing).

Stolen data vendors on the Open Web begin the sales process by first posting an advertisement about their product, specifying the product's price and the seller's preferred mode of communication (e.g., email, instant messaging, phone) (Franklin et al., 2007; Holt, 2013; Holt & Lampke, 2010; Holt, Smirnova, & Chua, 2016). The low cost of online advertising enables sellers to post multiple advertisements across various platforms to promote their products (Holt, 2013; Holt, Bossler, & Seigfried-Spellar, 2018; Holt & Lampke, 2010; Holt, Smirnova, & Chua, 2016). The level of information provided in advertisements vary across vendors, though sellers usually provide details related to product description, payment methods, modes of shipping, and refund/exchange policies (Holt, Bossler, & Seigfried-Spellar, 2018; Holt & Lampke, 2010; Holt, Smirnova, & Chua, 2016). In addition, sellers may also provide information about where and how they obtained their stolen data items (e.g., data sources), its various uses, and their years of experience as a vendor to enhance their perceived reputation as a reliable seller with quality product (Holt, Bossler, & Seigfried-Spellar, 2018; Holt & Lampke, 2010; Holt, Smirnova, & Chua, 2016).

Similar markets also operate on the Dark Web, a layer of the Internet that can only be accessed through specialized tools and search engines, such as The Onion Router (Tor) or the Invisible Internet Project (I2P) (Ablon, Libicki, & Golay, 2014; Barratt, 2012; Haasio, Harviainen, & Savolainen, 2020; Maimon & Louderback, 2019). Illicit online vendors on the Dark Web share many similarities with those on the Open Web, particularly in terms of the structure of their sites, advertisement content, and available products and services (Li & Chen, 2014; Smirnova & Holt, 2017).

Dark Web operations are unique in that anonymization tools such as Tor allow market participants to hide their identity and physical location by masking their computer’s Internet protocol (IP) address, rendering them unidentifiable and untraceable (Aldridge & Decary-Hetu, 2016; Barratt, Ferris, & Winstock, 2014; Dolliver, 2015; Lewman, 2016). Dark Web vendors are also less visible than Open Web vendors because specialized web browsers such as Tor are needed to access its content (Aldridge & Decary-Hetu, 2016; Barratt, Ferris, & Winstock, 2014; Dolliver, 2015; Holt, Bossler, & Seigfried-Spellar, 2018; Martin, 2014a, 2014b; Smirnova & Holt, 2017). Further, Dark Web vendors’ frequent use of cryptocurrency (e.g., Bitcoin) make tracing illicit economic transactions a challenging and laborious task (Aldridge & Decary-Hetu, 2016; Cox, 2016; Li & Chen, 2014; Smirnova & Holt, 2017).

The dual market environment adds to the complexity of decision-making for both buyers and sellers in the online stolen data marketplace as they must determine which platform best suits their operational needs. There are other issues that complicate this decision, creating the presence of information asymmetry, which is generally defined as imperfect information which lead to market failures (see Stiglitz, 2002; Tzanetakis et al., 2016). Information asymmetry occurs when different stakeholders possess conflicting information regarding a particular economic transaction or enterprise (see Connelly, Certo, Ireland, & Reutzel, 2011; Stiglitz, 2002). In stolen data markets, detailed material knowledge about vendors' motives and product qualities are concealed from prospective customers. This condition generates conflict between vendors who know their products' true quality and buyers who would make better informed decisions if they were privy to such information (Connelly et al., 2011; Stiglitz, 2002).

Since potential buyers are faced with copious amounts of vendor options and (mis)information that may lead to confusion, their ability to make sound decisions as to which vendors they can trust are limited (see Tzanetakis et al., 2016). Tzanetakis and colleagues (2016) argue that having a wide variety of vendor options and product information generates choice paralysis within buyers who are tasked with determining whom they should conduct business with (see also Barratt, Ferris, & Winstock, 2014). Choice paralysis, or the paradox of choice, is based on the idea that having an overabundance of options to choose from can cause individuals to stress and complicate decision-making, resulting in higher levels of anxiety, indecision, and dissatisfaction (see Schwartz, 2004; 2014). These complications force reliable vendors to generate effective signals of trust to procure profit, while prospective customers are required to look for signals to differentiate credible sellers and products from those dishonest and unreliable (Holt, Chua, & Smirnova, 2013).

Signaling theory complements rational choice perspectives by suggesting criminals deliberately calculate the costs and benefits associated with a signal before sending it (Gambetta, 2009). Those on the receiving end of signals make rational choices to either accept a signal at face value or interpret it as being deceptive (Gambetta, 2009; Decary-Hetu & Leppanen, 2016). Signals are used by participants in all manner of illicit markets on and offline, including to increase their trustworthiness and profit, evaluate others' reliability, and lower their risk of harm and misfortune. Examples of signals used by online data sellers include providing detailed product descriptions, such as specifying the card type (e.g., Visa Debit/Credit, Mastercard Gold) or product value (i.e., amount of available funds contained in the product); using more secure and encrypted payment methods (e.g., escrow, cryptocurrency); providing customers with free samples and product replacements; and presenting buyers with a space to post vendor ratings and product evaluations on dedicated customer feedback platforms (see Holt, Smirnova, & Hutchings, 2016).

Prospective buyers in online markets must identify and interpret vendors' signals as they were intended. Those who do so successfully should avoid being defrauded, while those who misinterpret signals will likely experience varying levels of financial loss (Connelly et al., 2011; Gambetta, 2009). This relates to the rational choice framework which argues that accurately assessing a behavior's costs and benefits can improve market performance (Decary-Hetu & Leppanen, 2016). Since the online illicit marketplace is an inherently untrustworthy platform, generating effective signals of trust and identifying credible actors from those unreliable are imperative to criminal success and longevity, especially since there are no formal avenues for dispute resolution when conflicts arise (Decary-Hetu & Leppanen, 2016; Holt, Smirnova, & Hutchings, 2016).

One potential signal that customers may immediately assess is the price vendors set for their products. Buyers are expected to assess whether a product is offered considerably lower or higher than its expected cost (Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013).

Research suggests markets operate on informal relational structures where the price for data is largely a function of how much people are willing to pay for it (Herley & Florencio, 2010; Holt, 2013; Holt, Chua, & Smirnova, 2013). This suggests it is not entirely clear whether pricing that is too low or too high poses a problem, as vendors who advertise stolen data products at either end of the pricing spectrum could be dishonest sellers looking to entice and defraud inexperienced buyers with inoperative products (Holt, Chua, & Smirnova, 2013).

To that end, Herley and Florencio (2010) suggested vendors in online stolen data markets use the price point for a given item as a signaling mechanism to express their credibility (see also see Akerlof, 1970; Holt, 2013; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010). Their argument is grounded in the idea that reliable sellers would be unwilling to offer functional products at substantially lower rates since it would be irrational to sell economically valued items (e.g., functional credit and debit card data) for considerably less than their value. Vendors who frequently advertise stolen data in large volumes at low prices could be either unskilled offenders or dishonest sellers looking to beguile inexperienced customers into buying inoperative products at enticingly low price points (Holt, Chua, & Smirnova, 2013). This framework suggests lower price points are likely to be signals of vendor unreliability, whereas higher prices are likely to be markers of vendor reliability (e.g., Herley & Florencio, 2010).

While this argument presents a logical inference regarding product price point and vendor reliability, it was made over a decade ago when data breaches were still somewhat novel and offenders had more ways of exploiting these stolen data products in both online and offline environments (see Huang & Ban, 2019; IBM, 2020; National Technology Security Coalition, 2020; Ponemon Institute, 2019; SelfKey, 2021; Verizon, 2020). For example, more secure card technologies such as the chip-and-PIN method, which requires the use of a personal identification number (PIN) to complete a transaction, had not been officially adopted by many countries at that time, including the United States (see Kagan, 2020). Stolen data markets have since progressed from a seller's market, where the available supply (e.g., less data breaches, vendors, and stolen data products) was disproportionate to the demand for product (e.g., more interested buyers), to a buyer's market where large volumes of stolen data products are regularly acquired, offered, and sold online (Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Chua, 2016; Smirnova & Holt, 2017). This shift in market condition calls into question whether price is still reflective of product quality and vendor reliability when there is an abundance of both stolen data vendors and potentially functional products for sale.

It is plausible that lower and higher price point may now be a more reflective measure of vendor unreliability (e.g., warning signs indicating vendor untrustworthiness) contrary to the argument proposed by Herley and Florencio (2010). Active sellers in the current marketplace may operate under a different business model where they prey on unexpecting buyers and look to procure large profits on one-time purchases that do not contain functional data. A smaller proportion of vendors may still operate to sell functional products at moderate prices to maintain a credible seller-buyer relationship. The conflicting signals generated from these shifting market dynamics would make it difficult to identify dependable markers of vendor reliability writ large, as either end of the pricing spectrum could indicate risk and vendor dishonesty. In the absence of purchasing data to establish its true quality relative to its asking price, inferences can still be made as to what product and vendor-level signals suggest vendor unreliability and potential reliability. Specifically, lower prices may be associated with vendor unreliability whereas higher prices could be associated with increased potential reliability.

Despite the constant flow of data breaches and increasing number of stolen data markets offering financial data, few have examined vendors' signaling practices and their ties to vendor unreliability as reflected in the pricing structure of online stolen data products (see Holt, Chua, & Smirnova, 2013 for exception). Specifically, limited research has examined the social dynamics of trust (e.g., signaling behaviors) that relate to the pricing structures of dumps products (e.g., digital information contained in active credit and bank card magnetic strips) advertised in online stolen data markets (see Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Chua, 2016). In addition, few have explicitly examined stolen data vendors' signaling practices in relation to product price point on both the Open and Dark Web, as well as across different types of vendors (e.g., shops and forums) within the same analysis (see Smirnova & Holt, 2017).

The lack of price structure analyses on stolen data products may be a result of the complexities involved in accurately measuring products' price point (see Holt, Smirnova, & Chua, 2016). For instance, vendors may offer stolen card data at different price points based on its country of origin (e.g., currency) and economic value (e.g., card/account balance) (Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Smirnova & Holt, 2017). Some vendors may also offer incentives such as bulk discounts to their already sophisticated pricing structures (Franklin et al., 2007; Holt & Lampke, 2010; Holt, Smirnova, & Chua, 2016). These complications may make disaggregating individual product prices a difficult endeavor and dissuade researchers from examining the economic structure of stolen data vendors and their products (see Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Chua, 2016; Smirnova & Holt, 2017 for exception).

Understanding the relationship between signaling behaviors and vendor unreliability is important as it could assist law enforcement devise effective intervention strategies targeted at disrupting the online illicit marketplace (see Holt, Blevins, & Kuhns, 2014; Holt, Chua, & Smirnova, 2013; Jacobs, 1996a; 1996b). For instance, identifying the signaling mechanisms that convey unreliability in the underground marketplace could help undercover law enforcement operations disguise themselves as potentially reliable sellers to both take down and apprehend serious market actors and participants (see Jacobs, 1996a; 1996b).

Relatedly, a deeper understanding of vendors' signaling behaviors at both the product and vendor-level would help investigative operations differentiate substantive mechanisms and practices from less concerning noise in terms of identifying serious market actors and warning signals from those that pose lower levels of threat (see Holt, Chua, & Smirnova, 2013). If a particular payment or communication platform can be identified as being a significant marker of potential reliability (e.g., less vendor unreliability), rendering that platform inoperative may induce an interruption large enough to disrupt the online illicit marketplace generally. This would save both law enforcement time and investigative resources, as well as generate a larger disruptive effect that obstructs illicit market behaviors. From a theoretical perspective, such research also extends understanding of traditional criminological theory by illustrating its ability to explain emerging forms of crime and deviance.

Given this gap in the literature, the current study explored the signaling mechanisms of 1,055 stolen data products across 40 vendors on the Open (n = 8; 20%) and Dark Web (n = 32; 80%). The current study used a signaling theory framework to examine whether vendors' differential use of product-level (e.g., detailed product information) and vendor-level indicators (e.g., payment methods, customer service features, customer feedback mechanisms) predicted stolen data price points. The primary objective is to test whether product and vendor-level signals of unreliability are associated with the variation in price point. That is, unreliable signals will predict a lower price point while reliable signals will predict a higher price point. If thus demonstrated, price point could be considered a proxy measure of vendor reliability. The research and policy implications of this study in understanding the behaviors of online illicit stolen data vendors are explored in detail.

CHAPTER 2: LITERATURE REVIEW

Extant literature has demonstrated that a thriving, global marketplace exists for the sale of sensitive financial information, harvested through data breaches and other forms of hacking (Franklin et al., 2007; Holt & Lampke, 2010; Motoyama et al., 2011; Smirnova & Holt, 2017; Thomas & Martin, 2006). There is particular emphasis on the sale of financial data, such as credit and debit card numbers, as this information can be used to commit various forms of economic crime, including credit card fraud and identity theft (Franklin et al., 2007; Holt, Bossler, & Seigfried-Spellar, 2018; Holt & Lampke, 2010; Motoyama et al., 2011). Limited research has examined how certain market behaviors and social dynamics influence stolen data vendors' perceived unreliability as reflected in the price point of individual products (see Holt, Chua, & Smirnova, 2013). As a result, there is uncertainty as to how various product and vendor-level indicators relate to the pricing structure of stolen data products offered within online markets hosted on both the Open and Dark Web (see Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Chua, 2016).

Research exploring the signaling factors associated with stolen data price points are important as it helps identify the mechanisms that predict both sellers' unreliability and potential reliability (see Decary-Hetu, Paquet-Clouston, & Aldridge, 2016; Holt, Chua, & Smirnova, 2013). Such research can prove useful to law enforcement by alerting them to the signaling behaviors most important to imitate and disregard during undercover operations. Similarly, the findings can inform agencies of the best platforms that facilitate market operations, including payment systems and email servers, so that they can be taken down to disrupt illicit market behaviors (see Jacobs, 1996a; 1996b).

Before examining whether various product and vendor-level indicators are significant predictors of stolen data price point, a discussion of hackers and hacking behaviors will be provided. Then, an explanation of the online stolen data marketplace and both the Open and Dark Web will be provided. The chapter will conclude with a discussion on rational choice theory, information asymmetry, and signaling theory, providing insight into how product and vendor-level dynamics act as significant predictors of both vendor unreliability and potential reliability (i.e., less unreliability) as reflected in their stolen data price point.

Hackers and Hacking Behavior

Hacking is generally defined as the act of modifying a system's software and/or hardware in ways that were neither proposed by the creator nor in line with the creator's intentions (Holt, 2020; Holt & Bossler, 2016). This definition does not focus on the criminal intent of the actor, but on the act itself (i.e., altering software and/or hardware). There are a diversity of hacking behaviors and motives, though criminal hackers are thought to represent only a minority subset of the entire hacker community (Holt, 2007; 2020; Steinmetz, 2015; 2017). In fact, numerous studies found hackers commonly engage in open-source software programming and computer hardware manipulation that are not criminally motivated (Coleman, 2012; 2013; Himanen, 2001; Holt, 2020; Levy, 1984; Soderberg, 2008; Steinmetz, 2015; Taylor, 1999; D. Thomas, 2002; J. Thomas, 2005; Turgeman-Goldschmidt, 2008; Warnick, 2004). Thus, hacking and hackers are not solely defined by criminality and deviant behavior, but other cultural aspects and factors.

The extant literature on hackers' motivations demonstrates that hackers are primarily interested in technology and advancing both hardware and software programing, often engaging in individual and collaborative projects to generate innovative ways to solve technical security problems (Coleman, 2012; 2013; Levy, 1984; Steinmetz, 2015; Taylor, 1999). This does not negate the fact that various hackers may still be motivated by politics (McKenzie, 1999; Meikle, 2002; Taylor, 2005), online thrill-seeking, trespassing, and innovating different types of malicious software (Jordan & Taylor, 1998; Steinmetz, Schaefer, & Green, 2017; Turgeman-Goldschmidt, 2008; Wall, 2007). Within these malicious acts, individuals may also be inspired by creativity, ambition, and the desire for technological development (Levy, 1984; Steinmetz, 2015).

The role of innovation and creativity is evident in the ways that hackers have moved to monetize and profit from their skills by engaging in a range of attacks including bank fraud, extortion, denial of service attacks, identity theft, and intellectual property fraud (Decary-Hetu & Dupont, 2013; Grabosky, 2006; Holt, 2020; Taylor, 1999; Wall, 2007). Hackers interested in monetizing stolen data products employ various techniques to obtain financial data, including infecting computers with malware and psychologically manipulating people into disclosing sensitive financial information using phishing schemes or spam emails (Furnell, 2002; Holt, 2020; Holt & Bossler, 2016; Huang & Brockman, 2010; Hutchings & Holt, 2017; James, 2005; Mitnick & Simon, 2002; Ngo et al., 2020; Peretti, 2009).

Acquiring financial data through malware involves the creation of tools needed to infiltrate computer systems and procure stolen financial data (Zhuge, Holz, Song, Guo, Han, & Zou, 2009). These tools can either be invented by a malware author or purchased through the illicit underground market in the event the individual does not want to invest their own time or skills to develop a tool (Decary-Hetu & Dupont, 2013; Decary-Hetu & Leppanen, 2016; Ianelli & Hackworth, 2005). A variety of methods can be used to infect computers and critical infrastructure, including drive-by downloads where malware is silently installed through a compromised website (Decary-Hetu & Leppanen, 2016). In turn, the device can be remotely accessed by the malware distributor to control the infected system and obtain sensitive financial information (Decary-Hetu & Leppanen, 2016).

Rise of Stolen Data Markets

The placement of financial data into online spaces has engendered numerous opportunities for hackers to steal valuable information, resulting in an excess of data beyond hackers' ability for personal use (Franklin et al., 2007; Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010; Holt et al., 2015; Holt, Smirnova, & Hutchings, 2016; Holz, Engelberth, & Freiling, 2009; Hutchings & Holt, 2015; Motoyama et al., 2011; Wehinger, 2011; Yip, Webber, & Shadbolt, 2013). As a result, various hackers sell these data to others interested in monetizing stolen financial information on the online illicit marketplace (Franklin et al., 2007; Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010; Holt et al., 2015; Holt, Smirnova, & Hutchings, 2016).

The recent growth of incidents targeting financial data can be attributed to changes in hackers' behaviors. Skilled hackers increasingly offer pre-packaged tools and kits to less technologically proficient individuals on online markets so they can successfully execute an attack to acquire financial data (Cooke, Jahanian, & McPherson, 2005; Decary-Hetu & Dupont, 2013; Hyslip & Holt, 2019). This enables non-technical individuals to enter the stolen data marketplace as sellers, creating a crowded and competitive economy where prospective buyers are exposed to more sellers and product options (see Holt, Chua, & Smirnova, 2013; Tzanetakis et al., 2016).

Online illicit markets typically operate as either a forum or a shop (Holt & Lampke, 2010; Li & Chen, 2014; Smirnova & Holt, 2017). Forums provide sellers with a platform to advertise their products and services to interested parties. Additionally, potential buyers can operate in an open space to inquire about and purchase goods (Copeland, Wallin, & Holt, 2020; Cunliffe, Martin, Decary-Hetu, & Aldridge, 2017; Li & Chen, 2014; Smirnova & Holt, 2017). Sellers typically provide information regarding their products' price point, preferred communication lines, favored methods of payment, and terms of service (Copeland et al., 2020; Cunliffe et al., 2017; Li & Chen, 2014; Holt & Lampke, 2010; Smirnova & Holt, 2017).

Forums follow an asynchronous two-way mode of communication where buyers and sellers openly interact with one another (e.g., buyers to buyers; buyers to sellers; sellers to sellers) using posts and threads. Most forums have administrators or operators who manage the daily operations of the platform. In certain cases, administrators also vet market participants and assign particular roles and titles to members based on their record and activity within the marketplace (Decary-Hetu & Laferriere, 2015; Holt & Lampke, 2010). Some forum operators may even assign third-party regulators with the task of moderating disputes and maintaining order on their page should conflicts arise (Holt, 2012; 2013; Martin, 2014a; 2014b).

Similar to eBay and Amazon, many online stolen data forums employ reputation systems where participants rate each other based on their experiences with one another (Decary-Hetu & Laferriere, 2015; Motoyama et al., 2011). These peer-based rating systems enhance the level of trustworthiness in the market and reduce risk between market participants (Decary-Hetu & Dupont, 2013). Motoyama and colleagues (2011) found that participants with higher public ratings were more likely to receive private messages about their products and services from prospective customers than those with lower ratings, suggesting a positive association between higher public ratings and greater profits (see Decary-Hetu & Dupont, 2013).

Shops, on the other hand, are spaces created by a single operator intending to sell goods and services to prospective customers (Martin, 2014a; 2014b; Smirnova & Holt, 2017). Shops are different from forums in that operations are run directly by the seller without any third-party intervention or oversight (Copeland et al., 2020). Shops follow a one-way mode of communication where sellers post information about their products or services on their shop webpages. Some shop vendors provide customers with an open space to rate the vendor and report their experiences, though customer feedback platforms are not consistently available across all shops (Copeland et al., 2020; Smirnova & Holt, 2017). Anything posted on these public customer feedback sections can be altered by the shop operator, including negative customer feedback to stop prospective customers from perceiving them as unreliable (Smirnova & Holt, 2017). As a result, shops generally limit the amount of information prospective customers have to evaluate a vendor's perceived credibility.

Open and Dark Web Markets

Understanding the general structure of online illicit markets is essential, as is recognizing the context in which they operate on the Open and Dark Web. The Open Web is described as the largely accessible part of the web that can be reached by common Internet browsers, such as Google Chrome, Firefox, and Internet Explorer, as well as through search engines which capture web content for archival purposes. The Dark Web is a smaller layer of the web that can only be accessed using certain browsers, such as The Onion Router (Tor) (Ablon et al., 2014; Barratt, 2012; Haasio et al., 2020; Maimon & Louderback, 2019).

The Dark Web was first introduced as a way to support the freedom of press and provide a space for open discussions without political pressure (Kang, Brown, & Kiesler, 2013). Subsequently, it has evolved into a platform where underground markets take advantage of its enhanced anonymity and encryption features. Various lucrative business operations take place on the Dark Web, including fraudulent investment schemes and illicit drug and stolen data sales

(Lee et al., 2019; Soska & Christin, 2015). In fact, Lee and colleagues (2019) found investment schemes (e.g., Ponzi fraud) to be the highest grossing illicit business operation on the Dark Web, with an estimated annual market value of $150 million. Drugs and stolen financial data are also popular business operations on the Dark Web, reaching an estimated annual market value of $14.4 million and $10 million respectively (Lee et al., 2019; Soska & Christin, 2015).

A unique feature of the Dark Web is its use of encryption techniques to hide participants' identity and location, allowing market participants to conceal both their IP address and the IP address of the servers hosting the vendor (Copeland et al., 2020; Decary-Hetu & Giommoni, 2017; Dingledine, Mathewson, & Syverson, 2004; Gehl, 2016). Anonymity is further enhanced by vendors' frequent use of cryptocurrency, including Bitcoin, Litecoin, Dash, and Ethereum, which serve as digital currency that hide the identity of those involved in a transaction (Morselli, Decary-Hetu, Paquet-Clouston, & Aldridge, 2017; Nakamoto, 2019). These novel payment methods make engaging in illicit transactions more convenient and safer for all involved parties (Barratt, Lenton, & Allen, 2013; Van Hout & Bingham, 2013a; 2013b).

Rational Choice, Information Asymmetry, and Signaling Theory

The availability of stolen financial information and prevalence of online data markets makes it imperative for researchers and practitioners to develop a greater understanding of the factors influencing the behaviors of its participants. Though this is a sophisticated and technologically-dependent form of crime, it is driven largely by the practices of humans whose decisions shape the overall market (see Leukfeldt & Holt, 2019). So-called classical theories are particularly insightful for criminological research examining illicit market operations and economically motivated actors. Classical perspectives view humans as rational actors that make conscious choices to engage in behavior after weighing its anticipated costs and benefits (Becker, 1968; Clarke, 1997; Clarke & Cornish, 1985; Holt & Dupont, 2019). Individuals are likely to engage in behavior that increases pleasure and decreases pain (Becker, 1968; Clarke, 1997; Clarke & Cornish, 1985; Holt, Bossler, & Seigfried-Spellar, 2018).

These frameworks are different from positivist perspectives that presume human behavior is a result of internal (e.g., biological, psychological) or external (e.g., social) forces beyond individuals' control (see Cullen, Agnew, & Wilcox, 2014; Holt, Bossler, & Seigfried-Spellar, 2018). Positivist theories of crime focus on individuals' personal and social characteristics as root sources of offending behavior. These theoretical differences lead to conflicting responses to crime and deviant behavior, such that classical theories focus on punishing the offender for their crimes while positivist perspectives focus on offender treatment, rehabilitation, and reform (Cullen, Agnew, & Wilcox, 2014).

Classical perspectives apply multiple interrelated theories based on their views of human decision-making, including deterrence and rational choice theory. Deterrence theory suggests individuals are discouraged from committing crime if they believe the anticipated punishment will be certain, swift, and proportionately severe (Decker & Kohfeld, 1985; Gibbs, 1975; Nagin & Pogarsky, 2001; Piquero et al., 2011; Pontell, 1978; Pratt et al., 2017). Certainty of punishment refers to how likely the individual will be caught and punished for the offense; swiftness of punishment refers to how quickly the punishment follows the criminal act, not the apprehension of the offender; and severity of punishment involves the intensity of the punishment relative to the harm caused by the crime (see Cullen, Agnew, & Wilcox, 2014; Decker & Kohfeld, 1985; Gibbs, 1975; Nagin & Pogarsky, 2001; Pontell, 1978).

Rational choice theory suggests people engage in criminal behavior if the expected benefits surpass the anticipated pains associated with the act (Becker, 1968; Clarke, 1997; Clarke

& Cornish, 1985; Cullen, Agnew, & Wilcox, 2014; Holt & Dupont, 2019). Rational choice perspectives suggest individuals consider situational factors such as their likelihood of detection (e.g., the presence and/or absence of guardians) and previous experiences (e.g., direct/personal and indirect/vicarious) with that particular behavior (Clarke, 1997; Gibbs, 1975). One's likelihood of detection influences the rational choice process such that riskier situations with higher levels of apprehension and punishment are likely to dissuade individuals' involvement in criminal behavior (Becker, 1968; Clarke, 1997; Gibbs, 1975).

Detection can come from both human subjects (e.g., security guards, parents, bystanders) and inanimate objects (e.g., security cameras, alarm systems) depending on the situation and behavior (Clarke, 1997; Cornish & Clarke, 2003; Reyns, 2010). Previous experiences with a particular behavior can similarly impact the rational choice process, as positive experiences that result in reward are likely to encourage repeat behavior whereas negative experiences involving apprehension are likely to discourage future participation (Clarke, 1997; Cornish & Clarke, 2003; Gibbs, 1975).

Informal risks that come from other offenders can also influence the decision-making calculus. Research has found that drug dealers are often targeted for robbery by other offenders because they tend to carry money and valuables on their person and have little legal recourse in the event they are harmed (Cross, 2000; Jacobs, 1996a; 1996b; 2000; Jacobs, Topalli, & Wright, 2000; Jacques & Wright, 2013; Knowles, 1999). In terms of stolen data market operations, buyers may receive invalid, non-functional, or simply no product from unreliable sellers after funds are transferred (see Herley & Florencio, 2010; Holt & Lampke, 2010; Holt et al., 2015; Holt, Smirnova, & Hutchings, 2016; Motoyama et al., 2011; Wehinger, 2011). Sellers can also cheat buyers by giving them smaller amounts or lower quality products than what was advertised and agreed upon (Holt et al., 2015). This is particularly salient in online markets, as buyers are unable to physically verify the quality of the product before completing the economic transaction (Herley & Florencio, 2010; Holt & Lampke, 2010; Holt et al., 2015; Holt, Smirnova, & Hutchings, 2016).

Research examining traditional illicit market operations using rational choice perspectives suggests market participants balance the expected benefits of the economic transaction (e.g., procuring profit, obtaining a hard to acquire product or a desired product/service for a lower cost) with its associated risks (e.g., formal and informal apprehension, receiving a defective product/service) before committing to the behavior (see Hamid, 1998; Holt, Blevins, & Kuhns, 2014; Jacobs, 1996a, 2000; Johnson, Dunlap, & Tourigny, 2000; Johnson & Natarajan, 1995; Knowles, 1999; Lupton, Wilson, May, Warburton, & Turnbull, 2002; McSweeney, Turnbull, & Hough, 2008; Topalli, Wright, & Fornango, 2002; VanNostrand & Tewksbury, 1999). Studies exploring traditional, offline drug market operations found that drug dealers carried lower amounts of product on their person to decrease both their criminal charge if caught and reduce their risk of theft by other criminals (see Jacobs, 2000). Relatedly, research examining offline prostitution offenses found both sex workers and their customers adopted overt groping measures to illustrate that neither party was affiliated with law enforcement (see Holt, Blevins, & Kuhns, 2014). These studies demonstrate the various risk factors that influence illicit market participants' decision-making process, underscoring the utility of rational choice perspectives in examining illicit criminal market behaviors writ large (see Holt, Blevins, & Kuhns, 2014; Holt & Dupont, 2019; Jacobs, 2000).

Although traditional illicit markets are similar to online markets in their shared focus around executing economic transactions involving risk and reward, there are notable differences in the presence and extent of information asymmetry complicating the decision-making process for online market participants relative to those operating offline (see Akerlof, 1970; Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Stiglitz, 2002). Information asymmetry occurs when market participants (e.g., buyers and sellers) possess varying levels of information about the quality of a product, making it harder for buyers to identify credible sellers and quality items (called "peaches") from dishonest sellers and defective products (called "lemons") (see Akerlof, 1970; Connelly et al., 2011; Stiglitz, 2002). As a result, prospective customers are only willing to pay somewhere between the true cost of the quality item and the lemon product (Akerlof, 1970). This market condition forces credible sellers to either accept a lower price point than what they are comfortable with or leave the marketplace entirely since they are unable to receive the full value for their quality product. This generates a lemon market where the majority of advertised products are nonfunctional (see Akerlof, 1970).

The complications created by information asymmetry force prospective buyers to look for signals to differentiate credible sellers from those untrustworthy (Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Hutchings, 2016; Tzanetakis et al., 2016). Credible sellers are similarly tasked with implementing effective signals that convey their reliability and trust (Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Hutchings, 2016; Tzanetakis et al., 2016). Gambetta's (2009) signaling theory extends the concept of information asymmetry by providing deeper insight into how criminal actors identify themselves in relation to each other and signal trust in unreliable environments (see Holt, Smirnova, & Hutchings, 2016). In general, a sign is defined as a visual display or physical indication of a certain message, whereas signals are the actual messages proposed by the sender (Decary-Hetu & Leppanen, 2016; Gambetta, 2009). Signs become signals when the sender displays them with an intended purpose (Decary-Hetu &

Leppanen, 2016; Gambetta, 2009). Gambetta's (2009) focus on physical signals translates to the online environment in the form of written information (e.g., language, ratings, feedback) and choice of visual displays (Layton, Watters, & Dazeley, 2010; Maruna, 2012; Tzanetakis et al., 2016). These signals help convert abstract qualities such as trust and reliability into visible indicators (see Gambetta, 2009; Layton, Watters, & Dazeley, 2010; Maruna, 2012).

Individuals in online illicit markets are constantly faced with the difficult task of evaluating the credibility and proposed meaning of signals (see Przepiorka, 2010). Those who interpret vendors' signals as intended will likely avoid being caught and defrauded, while those who do not will likely experience varying levels of harm (e.g., formal apprehension) and financial loss (Connelly et al., 2011; Gambetta, 2009). From a vendor's perspective, it is always beneficial to signal trustworthiness regardless of whether one is reliable, especially when information asymmetry is present (Gambetta, 2009; Holt, Smirnova, & Hutchings, 2016). This adds to the complication of differentiating reliable sellers from those untrustworthy, as dishonest actors are likely to imitate the signals used by their reliable counterparts to entice inexperienced and unsuspecting buyers (Gambetta, 2009).

There are certain signals that trustworthy actors use that are too costly for dishonest sellers to replicate (Holt, Smirnova, & Hutchings, 2016). According to Gambetta's (2009) signaling theory, reliable sellers are incentivized to generate signals that are both inexpensive to transmit yet costly to replicate. There are two types of signals that sellers use to express their intentions: weak signals and persuading signals (Gambetta, 2009). Weak signals are those that are inexpensive and easy for dishonest sellers to imitate. For instance, self-identifying as a trustworthy seller (i.e., writing in plain text that they are a reliable seller) is an inexpensive signal that can be accomplished by anyone with relative ease.

Persuading signals are those that are inexpensive for reliable sellers to generate but too costly for dishonest sellers to replicate (Gambetta, 2009). For example, a reliable seller may not view building a lasting presence and positive reputation on the marketplace a costly endeavor, especially since this is in their own best interests. On the other hand, a dishonest seller will find such a task incredibly costly and laborious given the time and effort needed to form a favorable reputation, which would be lost after engaging in a few fraudulent transactions (Decary-Hetu & Laferriere, 2015; Holt, Smirnova, & Hutchings, 2016; Motoyama et al., 2011).

The concurrent transmission of weak and persuading signals by both reliable and dishonest sellers generates confusion for buyers who seek to identify trustworthy sellers and products before engaging in a transaction (see Holt, Smirnova, & Hutchings, 2016). Three cost conditions (i.e., equilibrium condition, uninformative condition, intermediate condition) are generated from sellers' employment of signals, with greater information asymmetry creating both more confusion and a less informative cost condition (Gambetta, 2009; Holt, Smirnova, & Hutchings, 2016). An equilibrium condition occurs when there is a clear distinction between reliable sellers relative to untrustworthy sellers. Buyers are able to clearly separate reliable sellers from untrustworthy sellers based on the quality of their signals. For instance, mentioning the use of an inactive payment method (e.g., Liberty Reserve) could present an equilibrium condition where prospective buyers are able to clearly identify a seller as unreliable.

An uninformative condition occurs when buyers are unable to differentiate credible sellers and their signals from those untrustworthy. For example, the use of cryptocurrency could present an uninformative condition where prospective buyers are unable to differentiate sellers' reliability since most online illicit vendors operate using cryptocurrency (Aldridge & Decary-Hetu, 2016; Cox, 2016; Holt, Chua, & Smirnova, 2013; Lee et al., 2019; Li & Chen, 2014;

Smirnova & Holt, 2017). The wide use of cryptocurrency within the online illicit marketplace may obscure the intended meaning behind this signal, generating an uninformative condition where buyers are unable to differentiate credible sellers from those untrustworthy based solely on this practice.

Lastly, an intermediate condition occurs when dishonest sellers are able to successfully imitate several signals performed by trustworthy sellers, creating a situation where buyers are able to accurately interpret only a portion of signals based on their quality (Gambetta, 2009). Limited research suggests the online stolen data marketplace may resemble an intermediate condition where buyers are able to identify certain vendors' signals as intended, while others are more obscure and challenging to interpret (see Holt, Smirnova, & Hutchings, 2016).

To that end, Herley and Florencio (2010) suggested vendors use the price point of individual products as signals to express their reliability (see also Akerlof, 1970; Holt, 2013; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010). This proposition is based on the notion that reliable sellers offering quality product are not going to give away functional data at substantially lower prices since it has clear economic value (Herley & Florencio, 2010). Though credible sellers know they may never be able to receive the full value of their items as long as the marketplace is host to dishonest sellers, they are unwilling to accept a price that is considerably lower than both the product's contained economic value and the effort it took to acquire the item (see Akerlof, 1970; Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Hutchings, 2016). Under this logic, vendors who offer large volumes of product at low prices are likely to be dishonest sellers seeking to prey on inexperienced customers (Holt, Chua, & Smirnova, 2013). This argument established an economic framework that suggested lower price points are likely to be signals of vendor unreliability, whereas higher prices are likely to be indicators of vendor reliability (e.g., Herley & Florencio, 2010).

It is important to note, however, that numerous market dynamics have shifted since this argument was first presented (see Herley & Florencio, 2010). For one, the stolen data marketplace has expanded to include both a wider global audience and a more comprehensive product base, comprising of many stolen data products from all over the world (Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Chua, 2016; Smirnova & Holt, 2017). In addition, the tools used to procure stolen financial data have developed such that more advanced and intricate methods are now implemented, such as the use of point-of-sales (PoS) malware to steal unencrypted financial information from devices designed to complete economic transactions (National Technology Security Coalition, 2020; Ponemon Institute, 2019). The ways in which credit and debit cards are used have also evolved, as major card companies now employ more secure technologies that provide additional layers of protection when completing economic transactions (see Huang & Ban, 2019; IBM, 2020; Kagan, 2020; National Technology Security Coalition, 2020; Ponemon Institute, 2019; SelfKey, 2021; Verizon, 2020). These current market conditions cast doubt on whether price point still functions as a dependable signal of vendor reliability.

Given the conflicting signals generated from shifting market practices, it may be more difficult than before to identify dependable markers of vendor reliability, as both ends of the pricing spectrum can indicate risk and vendor dishonesty. In theory, since vendors set their own price points, higher prices should be an indicator of greater vendor reliability. However, the surplus of stolen data sellers and products may generate a ceiling of reliability such that while greater reliability is expected, it is not absolute (Herley & Florencio, 2010; Wehinger, 2011). The online stolen data marketplace may have moved from what was more an intermediate condition (i.e., buyers are able to accurately interpret only a portion of signals based on their quality) to a potentially uninformative condition by virtue of these changes (i.e., buyers are unable to differentiate credible sellers and their signals from those untrustworthy) (see Gambetta, 2009). As a result, it may be more appropriate to identify prominent markers of vendor unreliability and potential reliability (i.e., less unreliability) since pricing at either ends seemingly contain risk and uncertainty.

Vendors' Signaling Behaviors

Understanding the signaling behaviors associated with stolen data price points is important as it helps identify the product and vendor-level mechanisms that predict vendor unreliability, as well as potential reliability (see Decary-Hetu, Paquet-Clouston, & Aldridge, 2016; Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Hutchings, 2016). Research examining the signaling mechanisms that relate to the pricing structure of dumps advertised in online stolen data markets has been scant (see Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Chua, 2016). This may be a function of the information asymmetry that impacts both buyers' and sellers' behaviors (Franklin et al., 2007; Holt, Blevins, & Kuhns, 2014; Tzanetakis et al., 2016).

Previous research suggests products' price structure may be a prominent signaling behavior used by vendors to express their credibility (e.g., Herley & Florencio, 2010; Holt, 2013; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010; Motoyama et al., 2011). Dishonest sellers may intentionally list products at lower price points to attract buyers who hope to make a greater return on their small investment. Some buyers may also seek out products with higher price points because of the perception that they will obtain functional products (Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Hutchings, 2016). Products listed at considerably lower price points may be especially tempting for inexperienced buyers, especially when the market is saturated with vendor options and deceptive signals.

Given that the market is influenced by social forces designed to maximize reward and minimize risk for both buyers and sellers, certain vendor behaviors may act as signals that directly impact how a vendor and product is perceived (see Holt, Chua, & Smirnova, 2013). There is particular emphasis on the practices of vendors that should directly shape their perceived unreliability in the market, including providing/omitting detailed product descriptions, differential use of payment methods, customer service mechanisms, and customer feedback (Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Wehinger, 2011).

Limited work on the norms of the market suggests practicing clear communication and providing customer service are associated with greater potential reliability (Chu, Holt, & Ahn, 2010; Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Wehinger, 2011). Customer service involves the use of any contact lines that can be employed in the facilitation of a transaction, including providing email addresses, phone numbers, or instant messaging IDs. These signals could also be easily imitated by dishonest sellers, calling to question how much value they truly have as dependable signals relative to misleading indicators (see Holt, Smirnova, & Hutchings, 2016). Limited research has explored the impact of both product and vendor-level factors in assessing vendors' perceived credibility. Given that products (i.e., sales) are nested within vendors (i.e., markets), a failure to concurrently assess both product and vendor-level indicators ignores the reality that market participants encounter when interacting with one another. Further research examining these factors is needed to improve our understanding of the markers that signal vendor unreliability and potential reliability.

Detailed Product Description

There are numerous product-level details that vendors can include when advertising their stolen data products online. For one, vendors can provide information regarding the product's card type (e.g., Visa, Mastercard) or affiliation with a specific financial institution (e.g., Bank of America, J.P. Morgan). In fact, vendors can opt to include a generic description of the card type (e.g., credit card, debit card) or a specific payment network processor (e.g., Visa, Mastercard, American Express). This level of product information informs the prospective customer of the precise stolen data product, as some buyers may be interested in acquiring a particular type of card due to its widespread use and acceptance over others (e.g., Visa and Mastercard over Discover and American Express). While this level of detail may seemingly indicate potential reliability, it can easily be falsified by dishonest sellers seeking to deceive inexperienced buyers given the inexpensive and uncomplicated nature of falsely assigning card types and network institutions to product descriptions (Herley & Florencio, 2010). Therefore, it can be argued that sellers who do not provide product information related to card type/institution are more likely to have higher product price points, indicating greater potential reliability.

Another product-level indicator that may signal potential reliability is detailed cardholders' information. This could include providing any level of information related to the cardholder's identity, such as cardholders' first or last name, contact information/email address of record, city/state/province of residence, and zip/postal code. Though unrelated to a cardholders' personal identity, some vendors may even include the card's expiration date, if applicable. It can be argued that providing cardholders' information within a product's description is a more time-consuming endeavor than simply posting the advertisement as a generic dump product. Further, adding a higher amount of product detail can indicate sellers' familiarity with the item, suggesting increased potential reliability. While dishonest sellers seeking to deceive inexperienced buyers could provide fabricated cardholders' information, this type of product detail may be too labor-intensive for untrustworthy sellers to imitate (see Decary-Hetu & Laferriere, 2015; Holt, Smirnova, & Hutchings, 2016; Motoyama et al., 2011). The time and effort needed to indicate a card type (e.g., Visa Debit, Barclay's Mastercard) may be far less than what is required to post more detailed cardholders' information given the expanse of options within criteria such as cardholders' name, place of residence, and zip/postal code. As a result, sellers providing detailed cardholders' information may be more likely to have higher product price points given the added time and effort needed to generate such information (see Herley & Florencio, 2010).

Since data generated in breaches are derived from individuals residing in a specific city, state, and country, vendors may also decide to include an indication of the product's affiliated country within its product description. Specifying this type of information provides prospective customers with a more detailed sketch of the stolen data product, as some buyers may only be interested in acquiring product from a certain geographic place depending on their intended use. This level of product detail can be easily fabricated by dishonest sellers seeking to defraud unsuspecting buyers since it takes very little effort to suggest one has data from certain places (Herley & Florencio, 2010). As a result, it may be that sellers providing product information related to the data's country of origin are more likely to have lower product price points, suggesting vendor unreliability (see Herley & Florencio, 2010).

Specifying data as originating from the U.S. may similarly influence price. U.S. data has historically been low in cost, potentially given its abundance, widespread acceptance, and global use (Holt, Smirnova, & Chua, 2016). As a result, though labeling products as being from the

U.S. requires very little effort and can be easily falsified by dishonest sellers (i.e., vendors may sell nonfunctional U.S. data to turn a profit), its historically low cost could be another factor affecting price point that merits consideration (Holt, Smirnova, & Chua, 2016**)**. In this context, it may be that sellers providing U.S. data are more likely to have lower product price points (Holt, Smirnova, & Chua, 2016).

Another product-level indicator that may signal vendor unreliability is the estimated product value. This can include providing details regarding the available funds in a cardholder's account within the item's description. Similar to the other product-level indicators, product value can be listed in general terms using a range (e.g., card contains between $1,000-$3,000; maximum limit of $5,000) or more specifically using exact figures and amounts (Holt, Smirnova, & Chua, 2016). Specifying the funds available in an account provides prospective customers with a more detailed explanation of the stolen data product, allowing them to evaluate the product's perceived worth relative to its advertised cost (see Holt, Smirnova, & Chua, 2016). This indicator serves a similar function as providing the mileage of a vehicle in a used car advertisement in that it helps interested buyers evaluate the product's worth relative to its cost and functionality. Further, a vendor's ability to specify the amount of funds available in an account suggests they had interacted with it to check and determine whether it was active (Holt et al., 2015; Holt, Smirnova, & Hutchings, 2016). Though this information could be fabricated, it is a high risk move when the customer purchases the account and finds a lower amount than advertised (Decary-Hetu & Laferriere, 2015; Holt, Smirnova, & Hutchings, 2016; Motoyama et al., 2011). In this context, sellers omitting information related to the value of any account may be more likely to have lower product price points.

Payment Methods

Research suggests vendors' differential use of payment methods may signal unreliability (Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011). Studies have found that many sellers in the online illicit marketplace prefer cryptocurrency as their payment method of choice (see Ablon et al., 2014; Copeland et al., 2020; Cunliffe et al., 2017; Martin, 2014a; Smirnova & Holt, 2017). Bitcoin is the favored cryptocurrency on Dark Web markets generally, with more than 99.8% of identified cryptocurrency addresses being linked to Bitcoin (see Lee et al., 2019). Bitcoin is a decentralized digital currency that uses cryptographic algorithms to produce a ledger without a central point of authority (Nakamoto, 2019).

The absence of a central authority allows users to transfer currency over the network (e.g., Bitcoin network) without declaring their offline identity, using pseudonyms instead of the real identity of users like traditional banking systems or currencies (Lee et al., 2019). These characteristics make tracing Bitcoin transactions to particular individuals and behaviors a difficult and challenging endeavor (FBI, 2012; Lee et al., 2019). Services like CoinJoin and CoinShuffling can be used to further complicate the tracing process, as they allow users to blend Bitcoin payments from multiple users into a single transaction, making it seem like one user is responsible for all the addresses in the transaction (Lee et al., 2019; Maurer, Neudecker, & Florian, 2017; Ruffing, Moreno-Sanchez, & Kate, 2014).

Although cryptocurrencies provide market participants with a safer way of executing financial transactions, there may be a mixed relationship between its use and product price point (Franklin et al., 2007; Herley & Florencio, 2010; Holt & Lampke, 2010). On the one hand, the use of cryptocurrencies may be associated with lower product price point since they allow for an instantaneous transfer of funds between the buyer and seller (Herley & Florencio, 2010). Platforms that allow for an instant transferring of funds may be exploited by unreliable sellers who hope to obtain funds quickly, then either block or ignore customers' purchase order (Herley & Florencio, 2010). Alternatively, cryptocurrencies are less risky and more secure than other electronic payment methods. Given that online payment systems are historically the favored payment method within online markets, there is uncertainty as to what role and effect cryptocurrency payment systems have on product price point and vendor unreliability (see Holt, Chua, & Smirnova, 2013).

The use of other instantaneous online payment systems such as PayPal, Web Money, and Perfect Money, may be associated with lower product price point and vendor unreliability since they allow for a rapid transfer of funds between buyers and sellers without the added layer of anonymity and protection (Herley & Florencio, 2010). In addition, if most market participants on the online stolen data marketplace use Bitcoin and other cryptocurrency to complete transactions, not making that shift may be an indicator of vendor unreliability as it could indicate a seller's lack of familiarity with normative market practices (see Holt, Chua, & Smirnova, 2013; Lee et al., 2019). In fact, Holt and colleagues (2013) found certain instantaneous online payment systems (e.g., WebMoney) were associated with lower dumps price points. Further research is needed to better understand the variations in online payment methods and its impact on vendor unreliability and product price point. These findings would be especially invaluable to law enforcement who seek to investigate money laundering crimes associated with economic fraud (Holt, Chua, & Smirnova, 2013).

Vendors' use of escrow payment systems may also impact product price point (see Holt, Chua, & Smirnova, 2013). An escrow payment involves assigning a third-party as an intermediary to increase the likelihood of a successful transaction. Payment is sent from the buyer to the escrow agent and released to the seller only after the buyer confirms receipt of the product (Holt, 2013; Wehinger, 2011). Escrow services are intended to limit buyers' risk, as sellers only get paid when items are successfully delivered and confirmed as functional (Holt & Lampke, 2010; Holt et al., 2015; Holt, Smirnova, & Hutchings, 2016). Although percentages range depending on the escrow agent and transaction amount, agents tend to retain a portion of the transaction total for payment of their services (Holt & Lampke, 2010; Holt, Smirnova, & Hutchings, 2016).

The security system involved in operating an escrow agent is likely to increase vendors' perceived credibility as a trusted seller with quality product. In fact, Holt and colleagues (2013) found escrow payments were associated with higher product price points, increasing dumps' price point by 297 percent (see also Holt, 2013; Wehinger, 2011). Escrow services can also be easily falsified by dishonest sellers seeking to attract unsuspecting buyers, as they could easily claim in their advertisements that they employ escrow services when they do not use such methods (Holt, Smirnova, & Hutchings, 2016). Further research on the impact of escrow payments on product price point is needed to clarify its relationship as a marker of potential reliability.

Customer Service Mechanisms

Customer service mechanisms (e.g., customer service lines, free samples, product replacements) may also be significant markers of vendor unreliability as reflected in their product price point (Holt, Chua, & Smirnova, 2013). Research suggests operating a dedicated customer service line (e.g., emails, phone numbers, instant messaging IDs) may increase perceived trust between market participants, as these contact lines can be used in the facilitation of a transaction (Holt, 2013; Holt & Lampke, 2010; Holt, Smirnova, & Hutchings, 2016).

Customer service lines function as signals that illustrate vendors' interest and willingness to satisfy customer requests and orders. Although sellers may falsely list on their sites that they have a dedicated customer service line, research revealed buyers were more likely to interact with vendors who provided them with a dedicated customer support line than those who did not (Holt, 2013; Holt & Lampke, 2010; Holt, Smirnova, & Hutchings, 2016). In fact, vendors offering customer service lines were associated with higher dumps price points, suggesting a positive relationship between offering customer support lines and increased potential reliability (see Holt, Chua, & Smirnova, 2013).

Another customer service mechanism that may signal vendor unreliability is the use of free samples (Franklin et al., 2007; Holt & Lampke, 2010). This mechanism may be deployed to attract inexperienced customers, as sellers may offer valid products as bait only to send nonfunctional data after payment is received (Herley & Florencio, 2010; Wehinger, 2011). Providing prospective customers with free samples would further diminish reliable sellers' willingness to participate in the marketplace as it would lower their ability to capitalize from their stolen data products. As a result, it can be argued that sellers offering free samples are more likely to have lower product price points, suggesting vendor unreliability.

In addition, sellers may also offer free product replacements for defective items as long as the request is filed within a specified timeframe after purchase (Franklin et al., 2007; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010). While this customer service mechanism would suggest greater reliability on the part of vendors, it may negatively impact their profits through the free distribution of product (Herley & Florencio, 2010). Dishonest customers may also claim their purchased product was inoperative and request free product replacements even if their original order was functional. Vendors offering product replacements may be associated with lower product price points as it may be utilized to attract customers with the intent to defraud them (Herley & Florencio, 2010). In fact, research found product replacements to be associated with lower product price points, providing preliminary support for a “lemon” market where lower prices may be used by unreliable sellers as a tactic to lure inexperienced customers (Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013).

Customer Feedback

Providing buyers with a customer feedback platform is another mechanism that may be used to signal vendor unreliability (Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010; Wehinger, 2011). Customer feedback platforms give buyers the ability to directly affect the public reputation of vendors through the use of positive, negative, or neutral comments that express both their experiences with the vendor and the quality of their product(s) (Holt, 2013; Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011). Since transactions occur in private settings outside the online shop or forum, customer feedback sections serve as a risk avoidance tool for buyers. The information provided in public posts presents useful details to participants regarding vendors’ services, communication practices, quality of product, and overall ease of purchase (Copeland et al., 2020; Holt, 2013; Hutching & Holt, 2015).

In order to enhance their perceived reputation as a trustworthy seller, vendors encourage market participants to post publicly accessible customer feedback on their sites (Holt, 2013; Holt & Lampke, 2010; Holt et al., 2015; Holt, Smirnova, & Hutchings, 2016; Wehinger, 2011). Credible sellers are expected to receive positive feedback for selling quality products, often resulting in more frequent sales and satisfied buyers (Lusthaus, 2012). In contrast, transactions involving defective products and poor customer service are likely to result in negative customer feedback, affecting the seller’s reputation as a potentially unreliable vendor. To that end, positive customer feedback is likely an indicator of potential reliability whereas negative or neutral customer feedback should indicate vendor unreliability (Holt, 2013; Holt & Lampke, 2010; Holt, Smirnova, & Hutchings, 2016; Motoyama et al., 2011).

It is important to note that not all customer feedback is valued equally. Positive comments that are vague or lacking adequate description of the transaction process or product may not be of value to potential customers (Resnick, Kuwabara, Zeckhauser, & Friedman, 2000). While feedback sections give buyers the ability to voice their thoughts and concerns, prospective customers must carefully calculate the legitimacy of these posts since they may be fabricated by the vendor or misleading entirely (Holt, 2013; Holt & Lampke, 2010). It is possible for dishonest sellers to create fake accounts for the sole purpose of posting deceitful feedback on their sites, or negative feedback on their competitors’ sites (Holt, Smirnova, & Hutchings, 2016; Hutchings & Holt, 2015). Numerous forums have designated moderators responsible for policing such matters, though customer feedback posts on shops can be controlled and altered entirely by the vendor or operator. Relatedly, the volume of feedback reveals little about the vendor or product, as new vendors and products are likely to have fewer comments of any sentiment than older ones. The significance of customer feedback on vendors’ perceived reputation depends on both the quality and quantity of information provided in the comments (Dupont, Cote, Savine, & Decary-Hetu, 2016).

There are also various reliability issues with customer feedback, including a low incentive for customers to provide any type of rating (e.g., reporting bias) and difficulties in distinguishing fabricated evaluations from genuine ones (Dupont et al., 2016; Josang, Ismail, &

Boyd, 2007). As a result, overreliance on public customer feedback systems to assess vendor unreliability may lead to erroneous assumptions (see Dupont et al., 2016). Despite these uncertainties, studies have linked positive customer feedback to increased potential reliability and negative customer feedback (e.g., negative and/or neutral) to vendor unreliability (Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010; Holt, Smirnova, & Hutchings, 2016; Motoyama et al., 2011; Wehinger, 2011). This is reasonable given prospective buyers are more likely to interact with vendors who have fewer negative customer feedback because of their perceived reliability and lower potential for financial harm (Holt et al., 2015; Holt, Smirnova, & Hutchings, 2016).

Notwithstanding recent growths in research exploring the online stolen data marketplace, few have examined the product and vendor-level mechanisms that impact stolen data sellers' perceived unreliability as reflected in the price point of individual products (see Decary-Hetu, Paquet-Clouston, & Aldridge, 2016; Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Chua, 2016). Given this gap in the literature, the current study used a signaling theory framework to examine whether various detailed product descriptions, vendors' differential use of payment methods, customer service features, and customer feedback mechanisms predict vendor unreliability as reflected in their stolen data price points.

Several hypotheses were developed regarding whether product and vendor-level signals of reliability are associated with the variation in price point:

Hypothesis 1: Vendors send signals about credibility, such that unreliable vendors show certain signals of unreliability, whereas potentially reliable vendors show certain signals of increased reliability.

Hypothesis 2: Signals of unreliability and potential reliability affect products' price point, such that showing signals of unreliability will be associated with lower price point, and signals of potential reliability with higher price point.

Hypothesis 3: Lower price points indicate vendor unreliability whereas higher price points indicate increased potential reliability.

This dissertation will test the following ideas associated with the price point for stolen data products. First, it is expected that products with detailed cardholders' information, omitting product information regarding card type/institution, using more secure payment methods such as escrow, offering dedicated customer service lines, and the presence of positive customer feedback are likely to predict higher product price points, which should be a reflection of greater potential reliability. In contrast, listing information regarding the data's country of origin, omitting product value within the item's description, and selling U.S. data; accepting less secure digital payment systems; and offering free samples and product replacements, are expected to be warning signals predicting lower product price point and vendor unreliability. There is uncertainty as to the impact cryptocurrency payment systems have on product price point, as while they provide an online instantaneous transfer of funds, its widespread use and enhanced encryption makes it a less risky and more secure method of executing financial transactions (Franklin et al., 2007; Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010).

The data used in this dissertation comprised of 1,055 stolen data products across 40 vendors on the Open (n = 8; 20%) and Dark Web (n = 32; 80%). Data were derived from English language stolen data advertisements located on both the Open and Dark Web. Advertisements were manually coded using content analysis techniques to generate quantitative variables reflecting both product details and vendor behaviors (see Holt, 2010; Holt, Chua, & Smirnova, 2013; Holt & Dupont, 2019; Holt & Lampke, 2010; Holt, Smirnova, & Hutchings, 2016).

CHAPTER 3: RESEARCH METHODOLOGY

The aim of this chapter is to provide a detailed description of the data and analysis used in this dissertation. First, an explanation of how these data were collected will be provided. Then, a discussion highlighting the ethical concerns with using online data will be presented. The chapter will conclude by identifying the dependent and independent variables used in the study, followed by an outline of the analytical plan guiding the study's analysis.

Data Collection

Data were derived from stolen dumps advertisements located on both the Open and Dark Web. A dump is a term used to refer to stolen digital information contained in active bank and credit card magnetic strips, including credit/debit card numbers and card verification values (CVVs) (Franklin et al., 2007; Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010; Wehinger, 2011). Vendors' sites were identified using search protocols operated by Open and Dark Web browsers using keywords such as "dump full CVV sale." To supplement the limited results generated from Dark Web search engines, the study included vendors listed in indexes such as the Hidden Wiki and other cryptomarket listings to locate sellers that had been previously identified (see Copeland et al. 2020; Flamand & Decary-Hetu 2019).

Data collection ensued for 17-months (August 2018 to December 2019), generating a total sample of 45 dumps vendors on the Open (n = 11; 24.4%) and Dark Web (n = 34; 75.6%), consisting of 1,078 unique products. Approximately 2.1% of the full sample was eliminated from the analysis through listwise deletion due to missing data, resulting in a final sample of 1,055 products across 40 vendors on the Open (n = 8; 20%) and Dark Web (n = 32; 80%).

The current data set reflects a convenience sample of stolen data vendors on the Open and Dark Web. For one, only advertisements communicated in English were included in the present sample (see Franklin et al., 2007; Holt & Lampke, 2010). Though it is uncertain how many English-based stolen data vendors are likely active at any given time, studies suggest vendors operating in Cyrillic-based languages (e.g., Russian or Ukrainian) comprise a substantive portion of the online stolen data marketplace, which may be derived from a lack of deterrence from governments in those countries (see Ablon et al., 2014; Goncharov, 2012; Holt & Bossler, 2016; Holt, Smirnova, & Hutchings, 2016; Peretti, 2009; Smirnova & Holt, 2017). This suggests the current sample cannot be generalized to sites operating in Cyrillic-based languages. The current sample is still reflective of some aspects of the broader stolen data marketplace since a considerable proportion of stolen data market behaviors occur via English-language sites (Smirnova & Holt, 2017).

In addition, the current sample did not contain any invitation-only vendors, comprising solely of registration-only shops and forums (Holt, 2013). Invitation-only markets are unique in that they seek to admit only those whose behaviors demonstrate valued skills needed within that closed market community (Holt & Dupont, 2019). Studies have found that those rejected for membership in invitation-only vendors likely reflect qualities and behaviors that may either pose economic or real harm to the collective, whether because they are thought to be unscrupulous or an undercover operative (see Holt & Dupont, 2019). As a result, private vendors may exhibit different signaling behaviors due to higher levels of trust between participants, thus constituting a different economic model (Holt, 2013; Smirnova & Holt, 2017). Though these data may not be representative of the holistic practices of online stolen data vendors generally, it provides a sizeable sample of shops and forums hosted on both the Open and Dark Web.

All vendors in the final sample utilized a registration system to access its content. As a result, usernames for each vendor’s site were created for access purposes (see Holt, 2010; Holt, 2013; Holt, Chua, & Smirnova, 2013). Registration systems allow individuals to join and participate in market behaviors by registering an account with the vendor (e.g., username and password). These registration vendors are different from publicly accessible vendors in their added layer of protection and security from the general public (Holt, 2010; Holt, 2013; Holt, Chua, & Smirnova, 2013; Markham, 2011). Registration vendors should not be confused with invitation-only vendors who have enhanced member filtering systems that keep outsiders from accessing their marketplace (Holt, 2010; Holt, Chua, & Smirnova, 2013; Markham, 2011).

Data were compiled by saving all webpages from each vendor’s site as hypertext markup language (HTML) files without engaging with site operators in any discussion, including when registering to the site (see Holt, 2010; Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Chua, 2016; Holt et al., 2015; Holt, Smirnova, & Hutchings, 2016). Text and images from each website were manually coded using content analysis techniques to create quantitative variables suitable for analysis (see Holt, 2010; Holt, Chua, & Smirnova, 2013; Holt & Dupont, 2019; Holt & Lampke, 2010; Holt, Smirnova, & Hutchings, 2016). Specifically, information related to product details (e.g., type of item, advertised descriptions, product prices) and vendor behaviors (e.g., shipping methods, bulk discounts, primary form of contact) were collected. This method of web-based data collection is arguably more accurate, efficient, and useful than conventional self-report or interview methods given traditional methods’ susceptibility to various reporting and selection biases (see Barratt & Aldridge, 2016; Cunliffe et al., 2017; Latour, Jensen, Venturini, Grauwin, & Boullier, 2012; Savage & Burrows, 2007; Thelwall, 2009).

Ethical Concerns with Using Online Data

Online data sources present unique ethical concerns when used for research (see Holt & Bossler, 2016; Holt & Dupont, 2019). In particular, there has been debate over whether data generated from online sources should be considered public or private content given its unique characteristics (Aldridge & Decary-Hetu, 2014; Kitchin, 2003; Wilkinson & Thelwall, 2011). This issue is important because if online content is deemed private, obtaining participant consent would be necessary. Those who believe online data is public content argue participant consent is unnecessary because data were collected by natural observation (see Aldridge & Decary-Hetu, 2014; Kitchin, 2003; Wilkinson & Thelwall, 2011). Obtaining participants' consent is unnecessary as long as the subjects and creators of the online content remain anonymous, including IP addresses (see Wilkinson & Thelwall, 2011). Others suggest participant consent is needed only if direct quotations are utilized (see Eysenbach & Till, 2001).

To avoid any ethical complications concerning the nature of online data, the current study did not include any information that could be traced to an identifiable subject or IP address. Moreover, no direct quotes, participant names, or revealing information regarding how these marketplaces were accessed were included in this analysis (see Holt, 2013; Holt, Chua, & Smirnova, 2013; Markham, 2011). All data were anonymized for the quantitative analyses, thereby limiting any likelihood of identification of vendors, buyers, or those whose card data may have been compromised.

Dependent Variable

The dependent variable for this analysis is the price point for dumps. *"Dumps price point"* was measured as a continuous variable, with an average product cost of $346.25. It is worth nothing that products were advertised using various currencies (e.g., Euro, Pound Sterling, U.S. Dollar, and Bitcoin). To ensure consistency in value, all listed prices were converted from their original currency to the equivalent U.S. Dollar (USD) value based on 2019 exchange rate listings found on www.statista.com. The 2019 exchange rate for Euro, Pound Sterling, and

Bitcoin to USD was €1.12, £1.2772, and □7225.42917, respectively (see Statista, 2019). The converted price in USD was calculated by multiplying the appropriate exchange rate with the listed price of the product. The lowest costing product was a dump priced at $5, while the highest costing item was a set of five cloned credit cards priced at $14,450.86. Given the positive skewness in dumps' price point, the log price point was used for each product as positive skewness in the dependent variable can produce problems for significance testing in analyses (see Figures 1 and 2; Holt, Chua, & Smirnova, 2013; Olivier & Norberg, 2010).

Independent Variables

A series of five product-level (e.g., detailed product information) and seven vendor-level (e.g., payment methods, customer service mechanisms, customer feedback) variables were included to examine sellers' signaling behaviors. Detailed product information consisted of five binary variables reflecting various aspects of a product's advertised description: "*Card type/institution"* (yes = 0; no = 1); *"cardholders' information"* (no = 0; yes = 1); *"country of origin"* (no = 0; yes = 1); "*product value"* (yes = 0; no = 1); and "*U.S. data"* (no = 0; yes = 1). "*Card type/institution"* measured whether the product had a specific card type (e.g., Visa, Mastercard, American Express) or financial institution (e.g., J.P. Morgan, Bank of America, Barclay's) explicitly stated in its advertised description. "*Cardholders' information"* captured whether the product's description explicitly listed any detail related to the cardholder's identity, including cardholders' first or last name, city/state, and zip code. Relatedly, "*country of origin"* captured whether the product explicitly indicated which country the data was coming from, while "*product value"* measured whether the product's monetary value (e.g., card amount) was specifically listed within its item description. Lastly, "*U.S. data"* captured whether the stolen data was listed as originating from the U.S.

As stated earlier, this study hypothesizes that advertisements that do not contain card type/institution and have cardholders' information are expected to have higher product price points since they reveal more detailed information about the product and require additional steps from the seller to generate (see Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Wehinger, 2011). Products that contain information regarding the country of origin, are advertised as U.S. data, and do not have product value listed are expected to predict lower product price points due to the fewer steps involved in generating such detail and its ability to be easily imitated by dishonest sellers.

Three categories (e.g., payment methods, customer service mechanisms, customer feedback) consisting of seven independent variables were included at the vendor-level to examine sellers' signaling behaviors. Payment method consisted of three binary variables, indicating vendors' use of the following payment systems: "*Cryptocurrency"* (no = 0; yes = 1); "*escrow"* (no = 0; yes = 1); and "*instantaneous online payment systems"* (no = 0; yes = 1). "*Cryptocurrency"* consisted of all payment methods that used digital currency and strong cryptography, including Bitcoin, Bitcoin Cash, Litecoin, Dash, and Ethereum. "*Instantaneous online payment systems"* consisted of less secure electronic payment methods that provided quick monetary transactions between parties, including WebMoney, PerfectMoney, and Paypal.

Payment systems that are both safer and require additional steps from the seller, such as escrow, are expected to predict higher product price point, signaling potential reliability (see Herley & Florencio, 2010; Holt, 2013; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010; Wehinger, 2011). Relatedly, less secure digital payment systems that generate an immediate transfer of funds are expected to predict lower product price point, indicating vendor unreliability (see Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013). It is unclear whether using cryptocurrencies as payment is associated with higher or lower product price point, as while they function as instantaneous payment methods, they are both secured by advanced cryptography and widely used by market participants (see Holt, Chua, & Smirnova, 2013).

Customer service mechanisms were measured using three binary variables, including vendors' use of "*customer service lines"* (no = 0; yes = 1), "*free samples"* (no = 0; yes = 1), and "*product replacements"* (no = 0; yes = 1). Vendors' use of customer service lines are expected to predict higher product price points because they help promote trust between buyers and sellers (see Herley & Florencio, 2010; Holt, 2013; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010; Holt, Smirnova, & Hutchings, 2016; Wehinger, 2011). In contrast, vendors' use of free samples and product replacements are expected to predict lower product price points, as these mechanisms may be deployed to lure inexperienced customers to their business (see Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Wehinger, 2011).

"Positive customer feedback" was measured as a binary variable, indicating whether pre-existing customer feedback were positive (no = 0; yes = 1). Comments indicating a satisfactory outcome and experience with a product and/or vendor were coded as positive feedback (e.g., 'good data,' 'reasonable price,' and 'data worked'). Positive feedback is expected to predict higher product price point because it assists prospective customers identify sellers' practices and assess their reputation before engaging in a transaction, whereas negative comments are expected to predict lower price points suggesting vendor unreliability (see Holt, 2013; Holt & Lampke, 2010; Holt, Smirnova, & Hutchings, 2016; Motoyama et al., 2011).

Two control variables were also included in the analysis: web platform and vendor type. "*Web platform"* was measured as a dichotomous variable, indicating whether the seller was operating on the Open Web (0) or Dark Web (1). The majority of stolen dumps vendors and products were hosted on the Dark Web (vendors: n = 32, 80%; products: n = 683, 64.7%), with Open Web (vendors: n = 8, 20%; products: n = 372, 35.3%) comprising a smaller proportion of the sample. Similarly, “*vendor type”* was measured as a binary variable, indicating whether the seller operated on a forum (0) or shop (1). The majority of vendors and products operated on shops (vendors: n = 38, 95%; products: n = 993, 94.12%), with far fewer hosted on forums (vendors: n = 2, 5%; products: n = 62, 5.88%).

Analytical Plan

A mixed-effects multilevel regression was estimated to examine the significant associations between product-level and vendor-level predictors on the log price points for dumps. A mixed-effects multilevel model was conducted to account for the nested nature of the data (e.g., products nested within vendors). An intercepts-only model was initially estimated to determine the amount of variation in products across vendors (see Table 3). The intercepts-only model gives each vendor its own intercept that begins the slopes. The variation in vendors’ price when all other predictors are zero (e.g., intercept only) is interesting itself, as low variation would suggest all vendors are close to the mean level of price. The findings revealed an interclass correlation coefficient of 0.686, which indicated a high amount of variation in products’ price point (product-level/level-1) was explained by differences between vendors (vendor-level/level-2). The high interclass correlation coefficient suggests there is variation in products’ price point that need to be explained. This result suggests a multilevel model would provide a more comprehensive analysis over a single-level ordinary least squares (OLS) regression model.

Given the dichotomous nature of the independent variables, a random intercepts model without random slopes (e.g., random coefficients) was generated. Further, coefficients were exponentiated given the log-transformation of the dependent variable, allowing for percentage change interpretations. It is worth noting that "*free samples"* was excluded from the final analysis due to a high standard error (2.63) generated from inadequate variable distribution (see Table 1). Further, some variables within the intended model showed high correlations (r > .400), generating concern as to the presence of multicollinearity (see Table 2). In fact, "*web platform"* was removed from the final analysis due to its high variance of inflation (VIF) score of 5.671. There were no issues with multicollinearity after excluding "*web platform"* from the model, as no variance inflation factor (VIF) was higher than 3.848 (see Hair, Black, Babin, Anderson, & Tatham, 1998; see also Table 2). All analyses were conducted using the "magrittr" (see Bache & Wickham, 2016), "lme4" (see Bates, Machler, Bolker, & Walker, 2015), and "sjstats" (see Ludecke, 2021) packages in the R statistical software (R Studio–1.2.5042).

To explore whether some of the predicted relationships hold in different contexts, four additional mixed-effects multilevel logistic regression models were attempted using "*card type/institution*," "*product value*," "*positive customer feedback*," and "*U.S. data*" as outcome variables. However, given the lack of variation across many of the key predictors, these multilevel regression models encountered a condition referred to as total/complete separation, which occurs when a category or range of a predictor has only one value of the response (Heckman, 2016; Zeng & Zeng, 2019). This condition generates a lack of variation among the variables necessary to estimate a functional model. Statistical models with total/complete separation often display extremely high odds ratios, standard errors, and confidence intervals that are unintelligible (Heckman, 2016; Zeng & Zeng, 2019).

When total/complete separation is observed, one possible solution is to remove some of the variables that are seemingly generating this condition (Heckman, 2016; Zeng & Zeng, 2019).

However, regardless of the combination of variables removed, these exploratory models still exhibited regression coefficients that were overly high and uninterpretable. More concerning is the lack of sound rationale and ethical premise guiding this methodological solution. Adding and removing variables that lack adequate variation without a theoretical or conceptual justification point towards unethical data mining tactics that are exercised simply to convert nonsensical results into coherent findings. Given these data limitations, the additional multilevel models could not be estimated.

While crude in comparison to what the multilevel logistic regression models would generate, multiple three-by-three chi-square tests were conducted with the key variables of interest (see Tables 5-8). These chi-square tests examined the significant associations between key predictors of interest and the aforementioned variables (e.g., card type/institution, product value, positive customer feedback, U.S. data), while controlling for vendor type (e.g., shop v. forum). Since the intercepts-only model revealed a high interclass correlation coefficient, suggesting the presence of variation in products across vendors, controlling for vendor type was deemed appropriate and justifiable. While the results generated from these contingency tables are unable to present a sophisticated understanding of how these relationships may hold under different contexts, the crosstabs are able to highlight significant associations between the variables of interest (e.g., chi-square value) and the strength of these relationships (e.g., phi coefficient).

CHAPTER 4: RESULTS

This chapter provides a detailed explanation of the findings generated from both the mixed-effects multilevel regression model and the multiple chi-square analyses. First, the results derived from the intercepts-only model will be presented, which was estimated to determine the amount of variation in products' price point across vendors. Then, Model 1 of the mixed-effects multilevel regression model will be presented, which examined the five product-level predictors on the log price point for dumps. Next, Model 2 findings will be discussed, which examined both product and vendor-level predictors on the log price point for dumps. A discussion of the model fit indices (e.g., AIC, BIC) will be provided, which highlights the multilevel model as a more comprehensive analysis over a single-level ordinary least squares (OLS) regression given its ability to account for the nested nature of the data (e.g., products nested within vendors). The chapter will conclude with a presentation of the multiple three-by-three chi-square analyses, revealing significant associations between the variables of interest, the strength of these relationships, and their implications.

Findings

Before estimating a mixed-effects multilevel regression, an intercepts-only model was conducted to determine the amount of variation in products' price point across vendors (see Table 3). Specifically, the intercepts-only model served as the baseline model that indicated whether there was variation in vendors' price point when all other predictors are zero. A low variation (i.e., low interclass correlation coefficient) would suggest all the vendors are close to the mean level of price. To that end, the findings revealed an interclass correlation coefficient of 0.686, which indicated that approximately 69% of the variance in products' price point was between vendors, with the remainder of price point variability occurring within vendors. This suggests a high amount of variation in products' price point (product-level/level-1) was explained by differences between vendors (vendor-level/level-2). The intercept variance for vendors' price point was also significant ($\tau_{00} = 1.50$, p <. 001), suggesting significant variation in products' price point across vendors. These results point to the appropriateness of estimating a multilevel model given its ability to explain variation in products across vendors.

Model 1 provides the results of product-level predictors on the log price point for dumps. The analysis found that vendors who did not include product information related to card type/institution and country of origin for the dump was significantly associated with higher price points for stolen data. Not providing information related to card type/institution increased dumps' price point by 232 percent (B = 1.20, *t* = 8.45, p <.001), while indicating country of origin increased dumps' price point by 180 percent (B = 1.03, *t* = 8.96, p <.001)[1]. In contrast, not providing information related to product value and selling U.S. data was significantly associated with lower product price points. Specifically, not including the product value within an advertised description decreased dumps' price point by 92 percent (B = - 2.54, *t* = - 14.34, p <.001), while selling U.S. data decreased dumps' price point by 12 percent (B = - 0.13, *t* = - 2.19, p <.05). Detailed cardholders' information was not a significant predictor of the price point for dumps.

Model 2 provides the results of both product and vendor-level predictors on the log price point for dumps. The analysis found that not providing the card type/institution of the dump was significantly associated with higher stolen data price points. In particular, omitting the card type or financial institution from the product description increased dumps' price point by 376 percent

[1] The change in percentage for price point is derived from the logged dependent variable (i.e., "*log price point*")

(B = 1.56, *t* = 9.96, p <.001). In contrast, ads excluding an item's product value and selling U.S. data were significantly associated with lower price points. Ads without product value were associated with a 93 percent decrease in the price point of dumps (B = - 2.68, *t* = - 9.54, p <.001), while the sale of U.S. data was associated with a 20 percent decrease in the price point of dumps (B = - 0.22, *t* = - 3.86, p <.001). While significant in Model 1, indicating the country of origin was no longer significantly associated with an increase in dumps' price point when both product and vendor-level indicators were examined (see Table 4, Model 2). Similar to Model 1, detailed cardholders' information was not a significant predictor of dumps' price point in the full model.

In terms of vendor-level indicators, the analysis found that using cryptocurrency, offering product replacements, and having positive customer feedback were significantly associated with higher stolen data price points. Specifically, using cryptocurrency was associated with a 596 percent increase in price point (B = 1.94, *t* = 5.57, p <.001), while offering product replacements was associated with a 385 percent increase (B = 1.58, *t* = 9.96, p <.001), and positive customer feedback by 134 percent (B = 0.85, *t* = 3.44, p <.001).

Using instant online payment systems and offering customer service lines were significantly associated with lower dumps' price point. The use of instant online payment systems was associated with a 95 percent decrease in price (B = - 3.08, *t* = - 7.16, p <.001), while offering customer service lines was associated with a 90 percent decrease (B = - 2.28, *t* = - 6.96, p <.001). Contrary to the hypothesized relationship, escrow was not a significant predictor of the price point for dumps.

To determine which model best fit the outcome variable being examined, both Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) scores were estimated. Both criteria indicate measures of model fit, with lower values suggesting a better overall fit

(Kuha, 2004). The lower AIC and BIC scores in Model 2 reveals it had a better model fit than Model 1 in explaining the signaling mechanisms involved in assessing vendor unreliability (see Table 4). This suggests important nuances and details would have been unidentified if variations in mean price point across vendors were unaccounted for (e.g., conducting a single-level OLS regression). In fact, neglecting to account for the nested nature of the data would have yielded both different and imprecise results (e.g., effect sizes, significance levels).

In sum, the full model (see Table 4, Model 2) revealed four significant predictors of vendor unreliability: not providing product value, selling U.S. data, using instant online payment systems, and offering customer service lines. There were also four significant predictors of increased potential reliability: not providing card type/institution, using cryptocurrency, offering product replacements, and the presence of positive customer feedback relative to product price point. These findings suggest various product and vendor-level signals are used by online stolen data sellers, some of which are potentially more reliable than others. On a larger scale, this suggests the marketplace may be operating in a state of uninformative signals where buyers are unable to distinguish trustworthy sellers and their signals from those unreliable due to the perpetual circulation of mixed signals (see Gambetta, 2009).

To explore whether the predicted relationships in the multilevel model hold in different contexts, multiple three-by-three chi-square tests were conducted with key variables of interest (see Tables 5-8). These chi-square tests highlighted the significant associations between “card type/institution,” “product value,” “positive customer feedback,” and selling “U.S. data” with other key variables of interest while controlling for vendor type (e.g., shop v. forum). It is worth noting that the cardholder information, escrow, and instant online payment system variables were all unable to compute chi-square values and phi coefficients across the forum category due to them only having one value of the response category (see Tables 5-8, “Forum” column). This lack of variation demonstrates why the study was unable to estimate functional multilevel logistic regression models with any of the key variables of interest.

Table 5 provides the chi-square results of key variables and card type/institution. Various significant associations were revealed. For one, detailed cardholders’ information (χ^2(1df) = 27.527(p <.001); ϕ = .162) and country of origin (χ^2(1df) = 36.031(p <.001); ϕ = .185) had a significant, weak positive relationship with card type/institution, whereas selling U.S. data (χ^2(1df) = 83.352(p <.001); ϕ = -.281), instant online payment systems (χ^2(1df) = 30.289(p <.001); ϕ = -.169), and product replacements (χ^2(1df) = 22.456(p <.001); ϕ = -.146) had a significant, weak negative relationship with card type/institution. Further, the chi-square analyses found that product value (χ^2(1df) = 121.859(p <.001); ϕ = -.340) and cryptocurrency (χ^2(1df) = 252.127(p <.001); ϕ = -.489) had a significant, moderate negative relationship with card type/institution.

While these results do not model determinants of a relationship nor predict the likelihood of an outcome, they reveal statistically significant associations between variables and the strength of these relationships. Though numerous variables were found to have significant associations, vendors’ use of cryptocurrency appeared to have the strongest relationship with a product’s indication of card type or financial institution (ϕ = -.489).

Table 6 provides the chi-square results of key variables and product value. Apart from detailed cardholders’ information, all variables were significantly associated with product value. Specifically, cryptocurrency (χ^2(1df) = 100.792(p <.001); ϕ = .309), instant online payment systems (χ^2(1df) = 179.486(p <.001); ϕ = .412), and product replacements (χ^2(1df) = 218.142(p <.001); ϕ = .455) had a significant, moderate positive relationship with product value. Card type/institution (χ^2(1df) = 121.859(p <.001); ϕ = -.340), country of origin (χ^2(1df) = 186.753(p <.001); ϕ = -.421), and customer service lines (χ^2(1df) = 96.647(p <.001); ϕ = -.303) had a significant, moderate negative relationship with product value.

The chi-square analyses also demonstrated that escrow use (χ^2(1df) = 342.040(p <.001); ϕ = -.569) had a significant, strong negative relationship with product value, whereas selling U.S. data (χ^2(1df) = 65.324(p <.001); ϕ = .249) and positive customer feedback (χ^2(1df) = 12.958(p <.001); ϕ = .111) had a significant, weak positive relationship with product value. Although numerous variables revealed significant associations with an item's indication of product value, vendors' use of escrow appeared to have the strongest relationship (ϕ = -.569). In contrast to the chi-square results of card type/institution, stronger relationships were revealed when testing for significant associations between product value and the key variables of interest.

Table 7 provides the chi-square results of key variables and positive customer feedback. All variables were significantly associated with positive customer feedback except for card type/institution and selling U.S. data. Product value (χ^2(1df) = 12.958(p <.001); ϕ = .111), cryptocurrency (χ^2(1df) = 20.771(p <.001); ϕ = .140), and customer service lines (χ^2(1df) = 32.309(p <.001); ϕ = .175) all had a significant, weak positive relationship with positive customer feedback. Escrow (χ^2(1df) = 35.672(p <.001); ϕ = -.184) had a significant, weak negative relationship with positive customer feedback. The chi-square results also revealed product replacement (χ^2(1df) = 226.262(p <.001); ϕ = .463) had a significant, moderate positive relationship with positive customer feedback, while both detailed cardholders' information (χ^2(1df) = 147.991(p <.001); ϕ = -.375) and country of origin (χ^2(1df) = 142.681(p <.001); ϕ = -.368) had a significant, moderate negative relationship with positive customer feedback. The lone variable that exhibited a significant, strong relationship with positive customer feedback was the use of instant online payment systems (χ^2(1df) = 341.876(p <.001); ϕ = .569).

Table 8 provides the chi-square results of key variables and the sale of U.S. data. All included variables were significantly associated with the sale of U.S. data, except for detailed cardholders' information and positive customer feedback. Product value (χ^2(1df) = 65.324(p <.001); ϕ = .249), cryptocurrency (χ^2(1df) = 25.761(p <.001); ϕ = .156), and product replacements (χ^2(1df) = 10.350(p <.001); ϕ = .099) all had a significant, weak positive relationship with selling U.S. data, whereas card type/institution (χ^2(1df) = 85.352(p <.001); ϕ = -.281), escrow (χ^2(1df) = 46.824(p <.001); ϕ = -.211), and instant online payment systems (χ^2(1df) = 14.625(p <.001); ϕ = -.118) had a significant, weak negative relationship with selling U.S. data. The only variable that demonstrated a significant, moderate relationship with the sale of U.S. data was customer service lines (χ^2(1df) = 113.226(p <.001); ϕ = -.328).

CHAPTER 5: DISCUSSION AND CONCLUSION

Global data breaches affecting both corporations and individual consumers have surged considerably over the past two decades, generating an increasingly competitive and saturated stolen data marketplace that offers prospective customers a wide assortment of product choices (Higgins, 2014; Holt, Chua, & Smirnova, 2013; Huang & Ban, 2019; National Technology Security Coalition, 2020; Ponemon Institute, 2019; Seals, 2014; SelfKey, 2021; Tzanetakis et al., 2016; Verizon, 2020). Although online stolen data market behaviors involve varying levels of complex technology, the sale of private data are committed by people who operate under the same human decision-making processes and motivations as traditional offline crimes (see Leukfeldt & Holt, 2019). While studies exploring stolen data market behaviors have increased, few have examined vendors' signaling practices in relation to products' price point. Understanding the relationship between signaling behaviors and vendor unreliability is important as it could provide law enforcement with valuable tactical insight targeted at disrupting the online illicit marketplace.

Given the gap in the current literature, the present study used a signaling theory framework to examine whether various product and vendor-level factors predicted vendor unreliability as reflected in their stolen data price points. It was expected that products with detailed cardholders' information, omitting product information regarding card type/institution, using more secure payment methods, offering customer service lines, and positive customer feedback were likely to predict higher product price points, reflecting greater potential reliability. In contrast, information regarding the data's country of origin, omitting product value within the item's description, and selling U.S. data; accepting less secure digital payment systems; and offering free samples and product replacements, were expected to predict lower product price point and vendor unreliability (Franklin et al., 2007; Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010).

The aim of this chapter is to provide a detailed discussion of the study's implications and results. First, an overview of the study's main findings and its direct implications on criminological theory and online illicit market behaviors will be presented. Then, the individual implications derived from both the multilevel analysis and multiple chi-square analyses will be discussed. The chapter will conclude with a discussion on policy implications, followed by an outline of the study's limitations and directions for future research.

Implications of Main Findings

The current study used a signaling theory framework to examine whether the price for dumps was significantly associated with various signals of vendor unreliability, including various aspects of the dumps themselves and vendors' differential behaviors. The findings from the mixed-effects multilevel model revealed four significant predictors of vendor unreliability (e.g., not providing product value, selling U.S. data, using less secure digital payment systems, offering customer service lines) and four significant predictors of potential reliability (e.g., not providing card type/institution, using cryptocurrency, offering product replacement, positive customer feedback).

Overall, these findings suggest various product and vendor-level signals are employed by online stolen data sellers, with some being more dependable indicators of potential reliability than others. Specifically, these findings suggest the online stolen data market may have shifted from an intermediate condition where buyers are able to accurately interpret a portion of signals based on their quality, to a potentially uninformative condition rife with information asymmetry where buyers are no longer capable of differentiating credible sellers and their signals from those untrustworthy due to the constant dissemination of mixed signals (see Gambetta, 2009). Short of knowing an item’s true quality and its advertised price point, it is difficult to ascertain which signals are absolute markers of vendor reliability. As a result, the current analyses point to signals that convey unreliability and potential reliability since extant research has demonstrated both ends of the pricing spectrum may contain levels of risk and vendor dishonesty (see Herley & Florencio, 2010; Wehinger, 2011).

In terms of its direct implications on criminological theory, this study demonstrates the utility of signaling theory in understanding online illicit market behaviors, as well as the value of classical perspectives (e.g., rational choice) in understanding the decision-making processes of online offenders more generally. Specifically, this study’s application of signaling theory to online stolen data markets demonstrated that online market participants function on rational choice principles of risk and reward in similar way to offline offenders, finding that less labor-intensive and easy to replicate behaviors were associated with unreliable sellers as reflected in their lower prices, and more detailed and time-consuming indicators with potentially reliable vendors as reflected in their higher prices.

In accordance with Gambetta’s (2009) signaling theory framework, credible sellers operating in the online illicit marketplace are incentivized to develop persuading signals, which are described as gestures and indicators that are inexpensive to create yet costly to replicate. Though unreliable sellers are also encouraged to communicate persuading signals that convey greater perceived credibility, the various costs (e.g., monetary, time, labor) associated with generating these signals may be too demanding for dishonest sellers to pursue with any regularity (see Gambetta, 2009). As a result, unreliable sellers seeking to take advantage of unsuspecting or inexperienced buyers may be more likely to operate weak signals, which are described as expressions that are both inexpensive to generate and easy to imitate (see Decary-Hetu & Laferriere, 2015; Gambetta, 2009; Holt, Smirnova, & Hutchings, 2016; Motoyama et al., 2011).

The simultaneous dissemination of weak and persuading signals by both reliable and unreliable sellers increases confusion for buyers who must successfully differentiate trustworthy sellers from those unscrupulous to avoid financial risk and harm (see Holt, Smirnova, & Hutchings, 2016). Depending on the volume and quality of signals being transmitted, the market can resemble either an equilibrium condition (i.e., clear distinction between reliable sellers relative to untrustworthy sellers), intermediate condition (i.e., buyers are able to interpret only a portion of signals based on their quality), or uninformative condition (i.e., buyers are unable to differentiate credible sellers and their signals from those untrustworthy), with greater information asymmetry generating both more confusion and a less informative cost condition (Gambetta, 2009; Holt, Smirnova, & Hutchings, 2016).

The current findings suggest the online stolen data marketplace may have evolved from what was more an intermediate condition when the market was seemingly a seller's market (i.e., less vendor and product options), to a potentially uninformative condition as more vendor and product options flooded the economy. This shift in market condition is not unique to the online stolen data marketplace, as uninformative cost conditions generated from information asymmetry and mixed signals can occur in other forms of illicit markets as long as saturation of products and vendors are present. In fact, the current conditions exhibited in the stolen data marketplace may be reflective of what other illicit goods markets will experience if saturation is reached and signals' intended meanings become blurred.

Market saturation suggests large numbers of vendors are constantly transmitting a mixture of weak and persuading signals to entice prospective buyers, generating a high volume of information asymmetry that makes the process of differentiating credible sellers and their signals increasingly challenging. This condition may breed a level of risk that is exceedingly higher than the anticipated gains of operating within the illicit marketplace, potentially leading buyers to exit the marketplace given the imbalance of risk to reward (see Akerlof, 1970; Connelly et al., 2011; Stiglitz, 2002; Tzanetakis et al., 2016). If buyers are unable to clearly distinguish substantive signals of trust from illusory indicators due to the surge of mixed signals and blurred meanings, their expected benefit from participating in the illicit marketplace may be diminished.

According to classical theory principles, illicit online marketplace behaviors can be deterred by both increasing the anticipated risk and lowering the expected benefit of operating in the marketplace (see Becker, 1968; Clarke, 1997; Clarke & Cornish, 1985; Cullen, Agnew, & Wilcox, 2014). In theory, overwhelming the market economy with mixed signals would encourage prospective customers to leave the marketplace since their risk of being scammed and defrauded exceeds their ability to accurately interpret vendors' signals and acquire personal profit (see Akerlof, 1970; Connelly et al., 2011; Stiglitz, 2002; Tzanetakis et al., 2016). Potentially reliable vendors may similarly exit the marketplace given both their increased risk of apprehension and diminished profit margins from low sales and fewer buyers (see Connelly et al., 2011; Stiglitz, 2002; Tzanetakis et al., 2016). As a result, an oversaturated marketplace replete with information asymmetry may comprise mostly of dishonest sellers and "lemon" (e.g., low quality, nonfunctional) products given the departure of potentially reliable sellers (see Connelly et al., 2011; Stiglitz, 2002). The buyer population may similarly be limited to inexperienced or unsuspecting customers, as more competent buyers may have drifted from these markets into more secure, closed environments where enhanced membership filtering systems are practiced and enforced (see Holt & Dupont, 2019). This suggests more precarious and serious offenders may be active in more protected and closed market environments.

To that end, it is possible for sellers and buyers who leave the saturated, uninformative marketplace to migrate to closed or invitation-only platforms where the social dynamics of trust are less obscure, potentially resembling an equilibrium or intermediate condition as opposed to an uninformative condition. This would suggest a displacement of illicit market behaviors to other environments as opposed to its complete removal, highlighting the utility of restrictive deterrence frameworks in understanding how offenders change their behaviors to lower their overall risk of detection and apprehension in response to sanctions (see Gibbs, 1975; Jacobs, 1996a; Jacobs & Cherbonneau, 2014; Moeller, Copes, & Hochstetler, 2016). The restrictive deterrence framework provides a means to understand the decision-making process of offenders and their behaviors despite their perceived risks of harm (see Holt, Blevins, & Kuhns, 2014; Jacobs, 1996a; 1996b; Johnson, Dunlap, & Tourigny, 2000; Johnson & Natarajan 1995; Topalli, Wright, & Fornango, 2002). Alternatively, it is still possible for buyers and sellers who leave the uninformative marketplace (e.g., registration-only markets, open markets) to make a complete departure from the online illicit market as they may be denied entry into these closed or invitation-only platforms (see Holt & Dupont, 2019).

The implications of this study may be slightly different for traditional, offline markets as information asymmetry is intrinsically lower in offline spaces given buyers and sellers are able to meet in person and verify the quality of product (i.e., visual observation of actor and product) before completing a transaction. Despite differences in spatial affordances and levels of risk, the classical theory principles of increasing risk such that it exceeds potential benefits, can still be applied in offline markets to deter criminal behavior. Understanding how illicit actors operate in their respective environments can inform law enforcement of what behaviors and signals to imitate to disrupt underground market behaviors writ large.

Another implication that can be drawn from this analysis is that in untrustworthy environments where signals are employed to distinguish credibility and trust, making offenders doubt each other’s intentions may be an effective strategy for success. While apprehending individual actors is also an option, a more effective solution may be to let the uninformative market condition naturally draft members out of participation. By creating an unsafe business environment, willing participants would leave on their own given the enhanced costs relative to its expected rewards (see Akerlof, 1970; Connelly et al., 2011; Stiglitz, 2002). To that end, it may be beneficial to allocate law enforcement resources to exploring restricted markets and their social dynamics of trust as opposed to open or registration-only markets, since these markets seem to be unraveling on their own given the effects of market saturation and information asymmetry.

Overall, this study demonstrated that signaling theory, and classical perspectives writ large, can be applied to various forms of economic crimes and transactional behaviors that take place within inherently risky settings and untrustworthy environments (see Becker, 1968; Clarke, 1997; Clarke & Cornish, 1985; Cullen, Agnew, & Wilcox, 2014; Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt & Dupont, 2019; Holt, Smirnova, & Chua, 2016; Holt, Smirnova, & Hutchings, 2016; Hutchings & Holt, 2017). The current findings suggest classical theory principles have value in deterring online illicit marketplace behaviors, highlighting these frameworks’ utility in both explaining online crimes and the human decision-making process of online actors.

Implications of Product-Level Signals

In accordance with the study's hypotheses, the findings demonstrated that less labor-intensive and easy to replicate product details were associated with lower prices, whereas more detailed and time-consuming product descriptions were associated with higher prices. These findings corroborate Gambetta's (2009) conceptualization of weak and persuading signals in untrustworthy environments. Specifically, not providing information related to a product's value was significantly associated with lower product price points, whereas not providing the card type/institution was significantly associated with higher product price points. These findings are sensible since listing the product value of an item provides prospective buyers with a more detailed description of the data, allowing customers to evaluate the product's perceived worth relative to its advertised cost. Further, indicating the amount of funds available within a stolen data account suggests the vendor had both accessed and interacted with it to ensure it was active and functional (Holt et al., 2015; Holt, Smirnova, & Hutchings, 2016). In this context, choosing not to include the item's product value may be a warning signal suggesting a vendor's unfamiliarity with the product.

While an item's product value can be falsified and easily imitated by dishonest sellers, it resembles a persuading signal that is inexpensive to generate but costly when buyers purchase the account and find a lower amount than advertised (Decary-Hetu & Laferriere, 2015; Holt, Smirnova, & Hutchings, 2016; Motoyama et al., 2011). In contrast, though providing a product's card type/institution may indicate reliability, this level of information could be a weak signal that is easily constructed by dishonest sellers seeking to deceive inexperienced buyers given the unsophisticated nature of assigning general card types and network institutions to product descriptions (Herley & Florencio, 2010). This suggests items that do not explicitly list product value and those that have card type/institution details can be regarded as product-level warning signals indicative of nonfunctional data and vendor unreliability.

Independent of its historically low cost, vendors may also advertise nonfunctional U.S. data at low price points to bait unsuspecting or inexperienced customers to turn a profit. This suggests selling U.S. data at low prices may also be associated with vendors' unreliability. The high level of information asymmetry generated from mixed signals reinforces the stolen data marketplace as operating within an uninformative signaling condition whereby signals of unreliability and potential reliability are blurred and indistinguishable (see Gambetta, 2009). While further research exploring the impact of product-level indicators on vendors' perceived unreliability is needed, these findings provide general support for Gambetta's (2009) conceptualization of weak and persuading signals within the illicit online stolen data market (e.g., Holt, Smirnova, & Hutchings, 2016).

Implications of Payment Methods

While numerous studies examined the significant relationship between escrow, digital payment systems, and vendor credibility (see Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Hutchings, 2016), there was uncertainty as to the positive or negative impact cryptocurrency use had on product price point (Franklin et al., 2007; Herley & Florencio, 2010; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010). The use of cryptocurrencies may be associated with lower product price point since they allow for an immediate transfer of funds between the buyer and seller (Herley & Florencio, 2010). Payment systems that enable instant economic transactions may be used by unscrupulous sellers to obtain funds quickly, then neglect customers' purchase orders (Herley & Florencio, 2010).

On the other hand, cryptocurrencies employ advanced encryption technologies that make it more secure and less risky than other digital payment methods (Lee et al., 2019; Maurer, Neudecker, & Florian, 2017; Morselli et al., 2017; Nakamoto, 2019; Ruffing, Moreno-Sanchez, & Kate, 2014). This relationship is further complicated by its widespread use and popularity among both credible and dishonest sellers on the online illicit marketplace, creating an uninformative cost condition where buyers are unable to differentiate trustworthy sellers and their signals from those unreliable (see Ablon et al., 2014; Copeland et al., 2020; Cunliffe et al., 2017; Lee et al., 2019; Martin, 2014b; Smirnova & Holt, 2017). Given that digital payment systems are historically favored within online markets, there were questions as to what role and effect cryptocurrency payment systems had on product price point and vendor unreliability (see Holt, Chua, & Smirnova, 2013).

The current findings suggest vendors' use of cryptocurrency is significantly associated with higher product price points, indicating potential reliability. It is possible vendors can falsify their acceptance of cryptocurrency, though there seems to be a level of assurance that overrides the potential risks involved. This may be a result of cryptocurrency being an untraceable payment method for both the buyer and seller (FBI, 2012; Lee et al., 2019). Using cryptocurrency may also suggest one's familiarity with the trends of the marketplace and its common mannerisms (see Holt, Chua, & Smirnova, 2013; Lee et al., 2019). Though further research exploring different types of cryptocurrencies and its impact on perceived vendor unreliability would benefit comprehension surrounding its signaling intent, the current findings suggest the use of a widely employed, secure payment method such as cryptocurrency is likely to indicate greater vendor reliability.

Contrary to the proposed hypothesis, vendors' use of escrow was both not significant and negatively associated with stolen data price point despite its added layer of security and reduced level of buyers' risk (see Holt, 2013; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010; Holt et al., 2015; Holt, Smirnova, & Hutchings, 2016; Wehinger, 2011). This may be another instance where vendors claim to service escrow payments on their sites when they do not operationalize such methods (Holt, Smirnova, & Hutchings, 2016). Given the ease to which dishonest sellers can falsely claim to use escrow payments, this signaling mechanism may be a weak signal employed by unreliable sellers looking to attract unsuspecting buyers (Gambetta, 2009). This finding emphasizes the importance of differentiating vendors' actual use of a payment system against fabricated claims of its use (see Holt, Smirnova, & Hutchings, 2016). Short of completing a transaction with a particular vendor, obtaining insight around their alleged use of escrow may be increasingly difficult. Though further research on the impact of escrow payments on product price point is needed to clarify its relationship as a marker of vendor unreliability, the current findings suggest not all secure payment platforms function as indicators of potential reliability. These findings underscore the presence of conflicting signals and inconsistent messaging (e.g., informational asymmetry) that generate higher levels of risk and uncertainty within the stolen data marketplace.

The current results found that less secure instantaneous online payment systems were significantly associated with lower product price point, suggesting vendor unreliability in the expected direction. Since digital payment systems (e.g., PayPal, Web Money, and Perfect Money) allow for a quick transfer of funds between market participants without the added layer of anonymity and protection, it is unsurprising it was associated with lower product price point and vendor unreliability (Herley & Florencio, 2010). In fact, vendors who accept less secure forms of online payment may be seeking to exploit inexperienced buyers who are unknowledgeable of how to acquire and/or use safer payment methods (see also Holt, Chua, & Smirnova, 2013). While digital payment methods in general are more advantageous for the seller since payments are sent and received by the vendor first before products are delivered to the buyer, the current findings suggest the use of less secure instant online payment systems are significantly likely to be warning signals indicative of vendor unreliability (Holt, Chua, & Smirnova, 2013).

Implications of Customer Service Mechanisms

Contrary to its hypothesized relationship, providing a customer service line was significantly associated with lower product price point, suggesting vendor unreliability. Previous research found that operating a dedicated customer service line increased perceived trust between market participants, illustrating vendors' willingness to satisfy customer requests and orders (Holt, 2013; Holt & Lampke, 2010; Holt, Smirnova, & Hutchings, 2016). In fact, studies found offering customer service lines were associated with higher product price points, suggesting a positive relationship between offering customer support lines and increased potential reliability (see Holt, Chua, & Smirnova, 2013).

In spite of the aforementioned findings, the current results seem to suggest offering customer service lines function as weak signals that unreliable sellers transmit to attract unsuspecting customers (Gambetta, 2009). Similar to other vendor behaviors that operate as warning signals, claiming to have a customer service line must not be mistaken for its actual use and operation of one, as any seller can falsely advertise the presence of a customer service line regardless of its actually deployment (see Holt, Smirnova, & Hutchings, 2016). The oversaturation of stolen data vendors and products may have diluted this signaling mechanism such that its value as a dependable marker of potential reliability may have been reduced, creating an uninformative cost condition consisting of higher informative asymmetry and greater risk (Gambetta, 2009).

In contrast to customer service lines, offering product replacements was significantly associated with higher product price point, suggesting increased potential reliability. This finding was unexpected, as offering product replacements negatively impacts vendors' profit through the free distribution of product (Herley & Florencio, 2010). It seemed rational to hypothesize that offering free data would be a tactic used by dishonest vendors to beguile unsuspecting customers since credible vendors with quality product would be unwilling to give away functional data for free (Herley & Florencio, 2010). One possible explanation for this relationship could be the change in market conditions. It may be the case that once the marketplace expanded into a buyer's economy, credible vendors had to modify their behaviors to convey reliability and acquire customers. This shift may have pressured those potentially reliable to add customer incentives to assure their clientele they were in fact credible sellers. Even if offering product replacements come at sellers' own financial expense, building a trusting relationship with prospective buyers may function to maximize credible vendors' long-term benefit (see Becker, 1968; Clarke, 1997; Clarke & Cornish, 1985; Holt & Dupont, 2019). Further research on the impact of changing market dynamics on vendors' signaling behaviors is needed to develop a more comprehensive understanding of customer service mechanisms and its relations to vendor unreliability.

Implications of Customer Feedback

In accordance with the study's hypothesis, positive customer feedback was significantly associated with greater potential reliability (Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010; Wehinger, 2011). Customer feedback platforms give buyers the ability to directly affect the perceived reputation of vendors through the use of positive, negative, or neutral comments that express both their experiences with the vendor and the quality of product(s) (Holt, 2013; Holt & Lampke, 2010; Motoyama et al., 2011; Wehinger, 2011). Since transactions occur in private settings outside the online shop or forum, customer feedback sections serve as risk avoidance tools for buyers, providing market participants with useful public information regarding vendors' services, communication practices, quality of product, and overall ease of purchase (Copeland et al., 2020; Holt, 2013; Hutching & Holt, 2015). Reliable sellers are expected to receive positive customer feedback that reflect their perceived credibility, often resulting in more frequent sales and satisfied buyers, while transactions involving defective products and poor customer service are likely to result in negative customer feedback, affecting sellers' perceived reputation as an unreliable vendor (Lusthaus, 2012).

Though customer feedback platforms give buyers the ability to voice their thoughts and concerns in public spaces, prospective customers must carefully calculate the legitimacy of these posts since they may be fabricated by the vendor or misleading entirely (Holt, 2013; Holt & Lampke, 2010). It is possible for dishonest sellers to create fake accounts for the sole purpose of posting deceitful feedback on their sites, or negative feedback on their competitors' sites, which generates confusion as to the dependability of customer feedback as a persuasive signal (Gambetta, 2009; Holt, Smirnova, & Hutchings, 2016; Hutchings & Holt, 2015).

Despite the possibility for customer feedback to be misleading weak signals, the current study found positive customer feedback significantly predicts higher product price point, indicating increased potential reliability (see also Holt & Lampke, 2010; Motoyama et al., 2011). This is sensible given prospective buyers are more likely to interact with vendors who have greater positive customer feedback because of the lower potential for financial harm (Holt et al., 2015; Holt, Smirnova, & Hutchings, 2016).

Implications of Chi-Square Analyses

While crude in comparison to what the multilevel logistic regression models would generate, multiple three-by-three chi-square analyses were conducted with key variables of interest (e.g., card type/institution, product value, positive customer feedback, U.S. data). These four variables were chosen to see whether the predicted relationships in the multilevel model would remain constant across different contexts. Though predictions concerning direction of relationship and likelihood of outcome cannot be inferred based on these chi-square tests, they point to markers of vendor unreliability and potential reliability that future work can build upon.

The current findings revealed the presence of numerous significant associations across varying strengths among the key variables of interest. Among those that exhibited statistical significance, the payment method variables had the strongest individual relationships across many of the key variables. For instance, cryptocurrency had the strongest relationship with card type/institution; escrow had the strongest relationship with product value; and instant online payment systems had the strongest relationship with positive customer feedback. Their strength in association suggests payment method may be an avenue worth investigating further as a potentially robust marker of vendor unreliability.

Exploring other outcome variables related to signaling theory and vendor unreliability is needed to observe how well these predicted relationships hold in different contexts. If other outcome variables reveal similar findings and point to common markers of vendor unreliability, this would improve and strengthen our understanding of vendors' signals within both the illicit online marketplace and untrustworthy environments writ large.

Policy Implications

Understanding the relationship between signaling behaviors and vendor unreliability is important as it could assist law enforcement devise effective intervention strategies targeted at disrupting the online illicit marketplace (see Holt, Blevins, & Kuhns, 2014; Holt, Chua, & Smirnova, 2013; Jacobs, 1996a; 1996b). For instance, law enforcement could flood the marketplace with various deceptive posts and mixed signals to further complicate market participants' process of interpreting signals as intended, which is known as a sybil attack (Holt, Smirnova, & Hutchings, 2016). The use of sybil attacks would not only increase the volume of information asymmetry and complicate the identification of signal to noise, but also add to the current uninformative cost condition where buyers are unable to differentiate credible sellers and their signals from those untrustworthy (Decary-Hetu & Laferriere, 2015; Franklin et al., 2007; Holt, Smirnova, & Hutchings, 2016).

Successful sybil attacks may cause both buyers and sellers to leave the marketplace since they are unable to clearly identify dependable markers of trust, thereby increasing their perceived level of risk and harm. Further, effective sybil attacks may cause significant market disruptions without the need for law enforcement prosecutions or arrests (Holt, Smirnova, & Hutchings, 2016). This would save both investigative resources, as well as generate a larger disruptive effect that obstructs illicit market behaviors.

Sybil attacks, however, may only be effective for a short period of time on poorly managed markets (Franklin et al., 2007; Holt, Chua, & Smirnova, 2013; Holt, Smirnova, & Hutchings, 2016). Better managed markets may employ moderators who filter falsified posts and comments, reducing the impact of sybil attacks on market operations. Further, sybil attacks may encourage illicit actors to innovate greater risk detection mechanisms and signals to avoid law enforcement apprehension (Holt, Smirnova, & Hutchings, 2016). This dilemma suggests the need for law enforcement efforts to combine both traditional and novel methods to infiltrate and disrupt the online illicit marketplace (Holt, Smirnova, & Hutchings, 2016).

This is particularly salient when disrupting illicit market operations that occur in shops. Instead of populating forums with posts containing mixed signals, law enforcement can increase the volume of information asymmetry and obscure the identification of signals by designing fake shops that transmit mixed signals. Inundating the marketplace with fake shops and mixed signals reinforces an uninformative cost condition that generates high levels of confusion and risk that may cause both buyers and sellers to exit the marketplace (Decary-Hetu & Laferriere, 2015; Franklin et al., 2007; Gambetta, 2009; Holt, Smirnova, & Hutchings, 2016). Blurring the intended meaning of conventionally weak and persuading signals such that the potential risks of operating within the stolen data marketplace outweigh the expected benefits may provide enough doubt and uncertainty for market participants to discontinue their illicit behaviors entirely (Decary-Hetu & Laferriere, 2015; Franklin et al., 2007; Gambetta, 2009; Holt, Smirnova, & Hutchings, 2016).

The current findings can also be used to inform law enforcement of which signals to avoid making in order to present themselves as potentially reliable sellers during undercover sting operations (see Hutchings & Holt, 2017). This would be particularly useful in markets that have not yet reached saturation or those that are not in a state of uninformative signals (e.g., invitation-only markets). This would allow law enforcement to target key market actors and disrupt their active networks and perceived reliability (Decary-Hetu & Giommoni, 2017). Posting advertisements using signals that transmit increased potential reliability could be a useful mechanism to facilitate credibility within a market that has little to no feedback tool for users (see Copeland et al., 2020). Soska and Christin (2015) suggest a small proportion of vendors are responsible for a large portion of the sales. By targeting these individuals, law enforcement could force a large number of participants to find new suppliers and rebuild trust with new vendors, slowing down the market economy. In fact, previous research has shown that attacks targeting reputation systems could be used to destabilize online illicit markets and their activities (see Decary-Hetu & Giommoni, 2017; Decary-Hetu & Laferriere, 2015).

Law enforcement agencies can also target the facilitators (e.g., payment processors, email servers) of the illicit marketplace to cause a larger disruption in market behavior without apprehending individual actors. Taking down prominent payment platforms and communication facilitators that market participants, honest or otherwise, use to conduct business would not only induce a disruption that prevents actors from continuing their illicit behaviors, but also save law enforcement time and investigative resources (Holt, Smirnova, & Hutchings, 2016).

Limitations and Future Research Directions

Despite the study's novel findings and contributions to the literature, it is not void of limitations. In light of the recent crackdowns on illicit market vendors writ large, it is plausible the current findings no longer reflect the current state or condition of the marketplace (Hutchings & Holt, 2017; Macdonald & Frank, 2017). Many vendors operating on the Dark Web have changed their uniform resource locator (URL) addresses to alternate ones to avoid law enforcement detection (Copeland et al., 2020). Further, the ways in which market participants interpret and disseminate signals may have also changed, potentially causing a shift in business models as well. As a result, it is not entirely clear whether the aforementioned signals and their intended relationships apply to the whole of the broader online illicit marketplace. Future studies would benefit from exploring how these relationships and signaling mechanisms hold across different product markets. For instance, while the current study helps inform knowledge around illicit markets that exchange digital products (e.g., stolen card data), it would be interesting to explore whether these findings remain constant in other digital goods markets, as well as across physical goods vendors operating online (e.g., illicit firearms and identity document markets).

Relatedly, the price points noted in the current study may not be representative of the definitive pricing of these products in the current marketplace (see also Cunliffe et al., 2017). Since the data used in the present study was collected in 2019, its findings can only function as a snapshot of the product list and pricing structure in 2019. Both the list of products and their pricing could have changed in either direction since then, especially given the recent law enforcement crackdowns on illicit market vendors writ large (Demant, Munksgaard, Decary-Hetu, & Aldridge, 2018; Munksgaard, Demant, & Branwen, 2016).

Another limitation of the current study is its inability to generalize findings to the wider stolen data market population. It is important to note that the current sample consisted of mostly shops, with only 2 forums (n = 2, 5%; products: n = 62, 5.88%). This suggests the findings cannot be generalized to the entire community of stolen data vendors in existence. In addition, all the vendors in the present study were operated using English. Though it is uncertain how many English and non-English vendors are active at any given time, research suggests a sizeable number of stolen data vendors are operated using other common languages, including Russian. The inclusion of only English advertisements further restricts the study's findings to only those in that category.

Further, all vendors in the current sample were from registration-only platforms, limiting the study's overall scope and reach, as its findings cannot be generalized to closed or invitation-only vendors. It is worth noting that different types of stolen data markets may operate on distinct economic models (Herley & Florencio, 2010; Wehinger, 2011). Research suggests open markets are populated by unskilled actors while those with greater knowledge and experience engage in closed markets that are hidden from outsiders (Herley & Florencio, 2010; Wehinger, 2011). Open markets may feature invalid data at prices that do not reflect the economics of more hidden and insulated markets (Herley & Florencio, 2010; Wehinger, 2011). The presence of multiple markets suggests there may be a floor with less reputable vendors at the lowest end, a middle where a mixture of reputable and less reputable vendors operate, and possibly a ceiling involving both types of vendors, though there is uncertainty around this given the closed and restricted nature of these high-end markets (Herley & Florencio, 2010; Wehinger, 2011). Future studies would benefit from having a more comprehensive sample of vendors, platforms, and languages, as doing so would increase its overall generalizability and external validity.

Another limitation of the study is its inability to determine whether product advertisements were generated by actual sellers or other interested parties such as law enforcement or cybersecurity researchers (see Holt, Smirnova, & Hutchings, 2016). Many of the vendors contained in the present sample could have been fake sellers operated by government personnel, law enforcement agents, or cybersecurity researchers. While this level of information may be difficult to obtain using online data scraping methods, self-report surveys of online market participants would seem more effective (Aldridge & Decary-Hetu, 2016). Similarly, though the current study provided insight into the economic model of stolen data markets, it was unable to determine how many products were actually bought and sold since official transactions occur outside of public sight and observation (Franklin et al., 2007; Holt, 2013; Holt, Chua, & Smirnova, 2013; Holt & Lampke, 2010; Wehinger, 2011). A triangulated source of data collection would generate better economic models that can be assessed to evaluate victims' harm relative to offenders' return (see Aldridge & Decary-Hetu, 2016).

The current study also encountered various measurement and analysis issues. The most concerning measurement issue was the lack of variability among the predictor variables. For one, all the independent variables were dichotomous, limiting the ability to see differences across a range of options. A condition of total/complete separation was also present given the lack of variability across the key variables of interest. Regardless of how many distinct items a particular vendor offered, the variability at the vendor-level for those items were inadequate because they all came from the same seller. For instance, if a particular vendor accepted only cryptocurrency, all the items offered by this seller would also be associated with only cryptocurrency.

This generated a condition where some item categories had no counts for a given response option. It is worth noting that this data limitation may be a condition exhibited across many datasets exploring various illicit markets and products, given that products are nested within vendors. One possible solution to this problem could be to collect a large enough sample of vendors where total/complete separation does not occur. Alternatively, forum samples may have more variability across sellers since not every product offered on these forums are going to be from the same seller.

Relatedly, it may be beneficial to explore mixed methods research designs and interdisciplinary collaboration when studying online illicit market operations. Mixed research designs need to be introduced not only to identify the signals that work but also to understand exactly how criminals generate these signals and how they interpret them (Decary-Hetu & Leppanen, 2016). Research using both qualitative and quantitative approaches would provide us with greater insight into the criminal mindset and enable us to build more convincing signals, allowing law enforcement agencies to infiltrate criminal organizations more easily (Decary-Hetu & Leppanen, 2016). For instance, examining customer feedback sections using qualitative analyses to identify significant nuances regarding vendors' signaling behaviors could provide a more in-depth understanding of market participants' thoughts and behaviors (Decary-Hetu & Giommoni, 2017). Other methodological approaches such as conducting interviews and participant observation can further shed light on the signaling practices and rationales behind these illicit market actors. In terms of cross-disciplinary collaboration, integrating knowledge and techniques across computer science and criminology can help develop simulation models that provide better estimates of the impact of disruption attacks on online markets (e.g., sybil attacks) (Decary-Hetu & Laferriere, 2015).

Despite the study's limitations, the current findings provide knowledge around an underexamined area in both cybercrime and criminal justice research writ large. This study demonstrated a need for more robust datasets and multidisciplinary approaches to improve our understanding of illicit vendors' signaling behaviors relative to their perceived unreliability. Future research would benefit from extending this line of inquiry using more comprehensive samples and sophisticated methodological techniques to better understand the signaling behaviors influencing illicit online market operations.

APPENDIX

Table 1. Descriptive Statistics (N = 1,055)

Variable	Description	N	%	Mean	S.D.	Min.	Max.
Dependent Variable							
Price Point (USD)				346.25	1074.176	5	14450.86
Log Price Point (USD)				4.31	1.632	1.79	9.58
Product-Level Variables							
Card Type/Institution				0.23	0.418	0	1
	0=Yes	817	77.4				
	1=No	238	22.6				
Cardholders' Information				0.28	0.448	0	1
	0=No	762	72.2				
	1=Yes	293	27.8				
Country of Origin				0.53	0.499	0	1
	0=No	495	46.9				
	1=Yes	560	53.1				
Product Value				0.64	0.482	0	1
	0=Yes	385	36.5				
	1=No	670	63.5				
U.S. Data				0.44	0.497	0	1
	0=No	587	55.6				
	1=Yes	468	44.4				
Payment Methods							
Cryptocurrency				0.81	0.393	0	1
	0=No	201	19.1				
	1=Yes	854	80.9				
Escrow				0.19	0.389	0	1
	0=No	859	81.4				
	1=Yes	196	18.6				
Instant Online Payment Systems				0.23	0.420	0	1
	0=No	814	77.2				
	1=Yes	241	22.8				
Customer Service Mechanisms							
Customer Service Lines				0.82	0.387	0	1
	0=No	193	18.3				
	1=Yes	862	81.7				
Free Samples				0.02	0.152	0	1
	0=No	1030	97.6				
	1=Yes	25	2.4				
Product Replacements				0.43	0.496	0	1
	0=No	599	56.8				
	1=Yes	456	43.2				
Customer Feedback							
Positive Customer Feedback				0.27	0.443	0	1
	0=No	773	73.3				
	1=Yes	282	26.7				
Control Variables							
Web Platform				0.65	0.478	0	1
	0=Open Web	372	35.3				
	1=Dark Web	683	64.7				
Vendor Type				0.94	0.235	0	1
	0=Forum	62	5.9				
	1=Shop	993	94.1				

Table 2. Correlation Matrix of Model Variables

	PRICE POINT	TYPE	HOLD	COUNT	VALUE	U.S.	CRYPT	ESCRO	IOPS	CUSTOM	SAMPLE	PRODUCT	POSITIVE	SHOP
PRICE POINT	1													
TYPE	.356***	1												
HOLD	-.315***	.162***	1											
COUNT	.328***	.185***	.167***	1										
VALUE	-.689***	-.340***	.039	-.421***	1									
U.S.	-.262***	-.281***	.004	-.116***	.249***	1								
CRYPT	-.029	-.489***	-.481***	-.272***	.309***	.156***	1							
ESCRO	.504***	-.048	-.296***	.386***	-.569***	-.211***	.101***	1						
IOPS	-.191***	-.169***	-.337***	-.357***	.412***	-.118***	.264***	-.260***	1					
CUSTOM	.119***	.032	.031	.380***	-.303***	-.328***	-.117***	.144***	.257***	1				
SAMPLE	-.005	.065*	-.097**	.109***	.118***	-.089**	.076*	-.074*	.286***	.074*	1			
PRODUCT	-.269***	-.146***	-.127***	-.058	.455***	.099**	.068*	-.294***	.450***	.304***	-.136***	1		
POSITIVE	.062*	.028	-.375***	-.368***	.111***	-.039	.140***	-.184***	.569***	.175***	-.094**	.463***	1	
SHOP	-.236***	-.058	.155***	.064*	.012	.053	-.039	.119***	.136***	-.004	.039	-.099**	-.095**	1

Note. PRICE POINT = log price point; TYPE = card type/institution; HOLD = cardholders' information; COUNT = country of origin; VALUE = product value; U.S. = U.S. data; CRYPT = cryptocurrency; ESCRO = escrow; IOPS = instant online payment systems; CUSTOM = customer service lines; SAMPLE = free samples; PRODUCT = product replacements; POSITIVE = positive customer feedback; SHOP = vendor type.
*p<.05, **p<.01, ***p<.001

Table 3. Intercepts-only Model for Logged Price Point of Stolen Data (N = 1,055)

	Baseline Model			
Fixed Effects	β	exp (β)	SE	*t*
Constant	5.05***	156.02	0.20	24.80
Random Effects				
σ^2 (Residual-variance)	0.68			
τ_{00} (Intercept-variance)	1.50			
N (Vendors)	40			
Observations (Products)	1055			
AIC	2717.9			
BIC	2732.8			

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

Table 4. Mixed Effects Multilevel Regression for Logged Price Point of Stolen Data (N = 1,055)

	Model 1				Model 2			
Fixed Effects	β	exp (β)	SE	*t*	β	exp (β)	SE	*t*
Constant	5.12***	167.34	0.26	19.35	5.83**	340.36	1.88	3.10
Card Type	1.20***	3.32	0.14	8.45	1.56***	4.76	0.16	9.96
Cardholders' Information	-0.10	0.90	0.09	-1.11	0.06	1.06	0.08	0.75
Country	1.03***	2.80	0.12	8.96	0.06	1.06	0.15	0.40
Product Value	-2.54***	0.08	0.18	-14.34	-2.68***	0.07	0.28	-9.54
U.S. Data	-0.13*	0.88	0.06	-2.19	-0.22***	0.80	0.06	-3.86
Crypto					1.94***	6.96	0.35	5.57
Escrow					-1.33	0.26	1.05	-1.26
Instant Online Payment Systems					-3.08***	0.05	0.43	-7.16
Customer Service Lines					-2.28***	0.10	0.33	-6.96
Product Replacements					1.58***	4.85	0.16	9.96
Customer Feedback					0.85***	2.34	0.25	3.44
Vendor Type					-0.50	0.61	1.91	-0.26
Random Effects								
σ^2 (Residual-variance)	0.53				0.44			
τ_{00} (Intercept-variance)	2.10				6.80			
N (Vendors)	40				40			
Observations (Products)	1055				1055			
AIC	2483.0				2338.1			
BIC	2522.7				2412.5			

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

Table 5. Three-way Contingency Table for Card Type/Institutions, Controlling for Vendor Type (N = 1,055)

Variables	**Forum**		**Shop**		**Total**	
	χ^2	ϕ	χ^2	ϕ	χ^2	ϕ
Cardholders’ Information	–	–	32.045***	.180	27.527***	.162
Country of Origin	15.308***	-.497	55.746***	.237	36.031***	.185
Product Value	0.944	.123	138.152***	-.373	121.859***	-.340
U.S. Data	15.122***	-.494	69.584***	-.265	83.352***	-.281
Cryptocurrency	19.289***	-.558	238.007***	-.490	252.127***	-.489
Escrow	–	–	1.833	-.043	2.421	-.048
Instant Online Payment	–	–	28.605***	-.170	30.289***	-.169
Customer Service Lines	21.051***	-.583	5.612*	.075	1.114	.032
Product Replacements	50.062***	-.899	10.186**	-.101	22.456***	-.146
Positive Customer Feedback	1.575	.159	0.125	.011	0.803	.028

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

Note. All df = 1; Some values are missing due to low/no counts (e.g., total separation).

Table 6. Three-way Contingency Table for Product Value, Controlling for Vendor Type (N = 1,055)

Variables	**Forum**		**Shop**		**Total**	
	χ^2	ϕ	χ^2	ϕ	χ^2	ϕ
Card Type/Institution	0.944	.123	138.152***	-.373	121.859***	-.340
Cardholders’ Information	–	–	1.518	.039	1.624	.039
Country of Origin	26.458***	.653	237.729***	-.489	186.753***	-.421
U.S. Data	14.329***	-.481	84.947***	.292	65.324***	.249
Cryptocurrency	5.099*	.287	96.219***	.311	100.792***	.309
Escrow	–	–	349.225***	-.593	342.040***	-.569
Instant Online Payment	–	–	181.777***	.428	179.486***	.412
Customer Service Lines	2.375	-.196	95.015***	-.309	96.647***	-.303
Product Replacements	0.003	.007	234.650***	.486	218.142***	.455
Positive Customer Feedback	30.770***	-.704	29.071***	.171	12.958***	.111

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

Note. All df = 1; Some values are missing due to low/no counts (e.g., total separation).

Table 7. Three-way Contingency Table for Positive Customer Feedback, Controlling for Vendor Type (N = 1,055)

Variables	**Forum**		**Shop**		**Total**	
	χ^2	ϕ	χ^2	ϕ	χ^2	ϕ
Card Type/Institution	1.575	.159	0.125	.011	0.803	.028
Cardholders' Information	–	–	143.616***	-.380	147.991***	-.375
Country of Origin	32.317***	-.722	114.358***	-.339	142.681***	-.368
Product Value	30.770***	-.704	29.071***	.171	12.958***	.111
U.S. Data	10.041**	.402	4.053*	-.064	1.622	-.039
Cryptocurrency	1.342	-.147	23.775***	.155	20.771***	.140
Escrow	–	–	32.698***	-.181	35.672***	-.184
Instant Online Payment	–	–	373.833***	.614	341.876***	.569
Customer Service Lines	1.441	.152	31.172***	.177	32.309***	.175
Product Replacements	1.107	-.134	247.613***	.499	226.262***	.463

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

Note. All df = 1; Some values are missing due to low/no counts (e.g., total separation).

Table 8. Three-way Contingency Table for U.S. Data, Controlling for Vendor Type (N = 1,055)

Variables	**Forum**		**Shop**		**Total**	
	χ^2	ϕ	χ^2	ϕ	χ^2	ϕ
Card Type/Institution	15.122***	-.494	69.584***	-.265	83.352***	-.281
Cardholders' Information	–	–	0.016	-.004	0.020	.004
Product Value	14.329***	-.481	84.947***	.292	65.324***	.249
Cryptocurrency	4.705*	.275	23.197***	.153	25.761***	.156
Escrow	–	–	50.241***	-.225	46.824***	-.211
Instant Online Payment	–	–	16.724***	-.130	14.625***	-.118
Customer Service Lines	6.849**	.332	133.460***	-.367	113.226***	-.328
Product Replacements	18.728***	.550	6.213*	.079	10.350**	.099
Positive Customer Feedback	10.041**	.402	4.053*	-.064	1.622	-.039

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

Note. All df = 1; Some values are missing due to low/no counts (e.g., total separation).

Figure 1. Histogram of Price Point

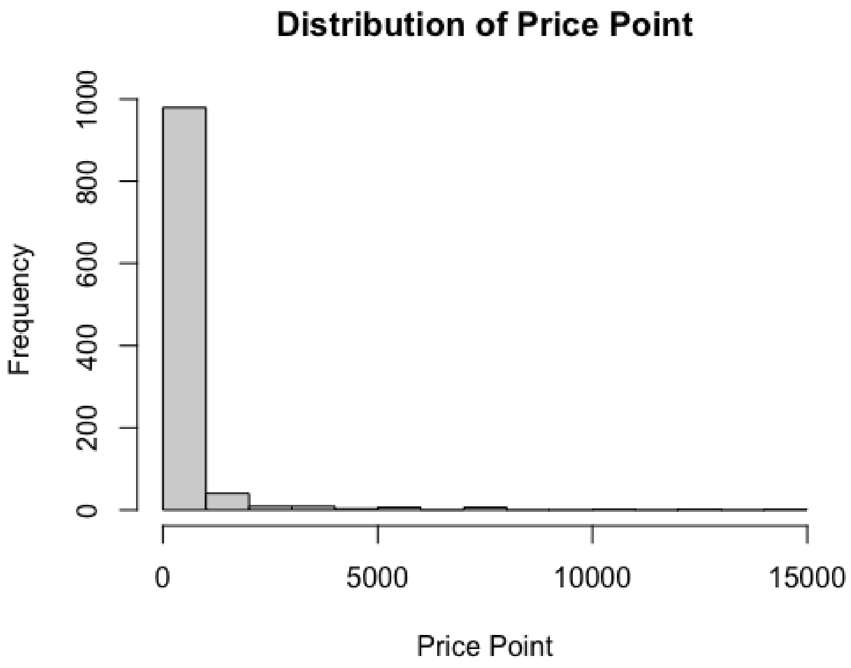
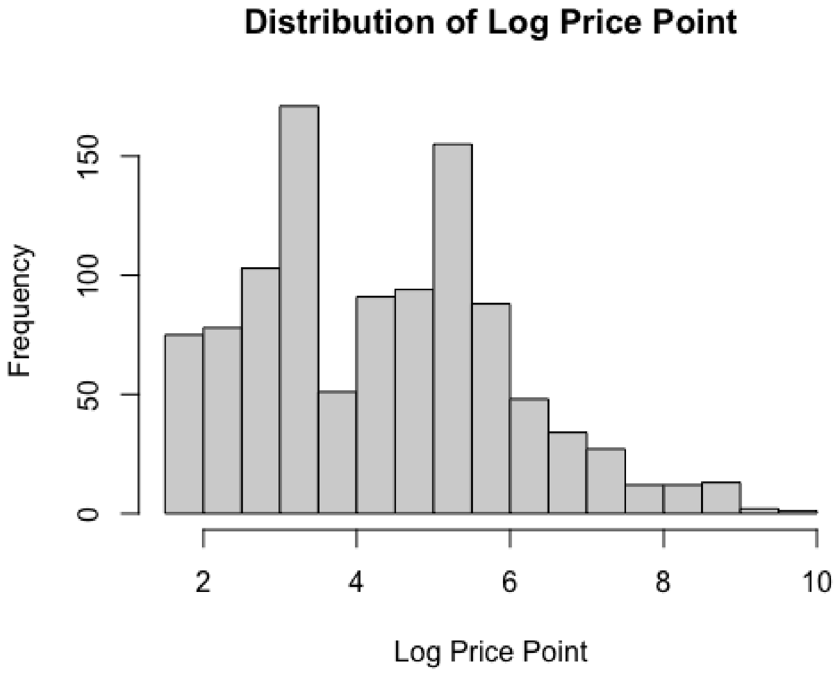


Figure 2. Histogram of Log Price Point



BIBLIOGRAPHY

BIBLIOGRAPHY

Ablon, L., Libicki, M. C., & Golay, A. A. (2014). *Markets for cybercrime tools and stolen data: Hackers' bazaar*. Santa Monica, CA: RAND Corporation.

Akerlof, G. A. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, *84*(3), 488-500. https://doi.org/10.2307/1879431

Aldridge, J., & Décary-Hétu, D. (2014). Not an 'Ebay for Drugs': The cryptomarket ‘Silk Road’ as a paradigm shifting criminal innovation. *Social Science Research Network.* http://dx.doi.org/10.2139/ssrn.2436643

Aldridge, J., & Décary-Hétu, D. (2016). Hidden wholesale: The drug diffusing capacity of online drug cryptomarkets. *International Journal of Drug Policy*, *35*, 7-15. https://doi.org/10.1016/j.drugpo.2016.04.020

Anderson, R., Barton, C., Böhme, R., Clayton, R., Van Eeten, M. J., Levi, M., Moore, T., & Savage, S. (2013). Measuring the cost of cybercrime. In R. Bohme (Eds.), *The Economics of Information Security and Privacy* (pp. 265-300). Berlin/Heidelberg: Springer. https://doi.org/10.1007/978-3-642-39498-0

Bache, S. M., & Wickham, H. (2016). Package ‘magrittr’. Retrieved from https://cran.r-project.org/web/packages/magrittr/magrittr.pdf

Barratt, M. J. (2012). Silk Road: Ebay for drugs. *Addiction*, *107*(3), 683. https://doi.org/10.1111/j.1360-0443.2011.03709.x

Barratt, M. J., & Aldridge, J. (2016). Everything you always wanted to know about drug cryptomarkets*(* but were afraid to ask). *International Journal on Drug Policy*, *35*, 1-6. https://doi.org/10.1016/j.drugpo.2016.07.005

Barratt, M. J., Ferris, J. A., & Winstock, A. R. (2014). Use of Silk Road, the online drug marketplace, in the United Kingdom, Australia and the United States. *Addiction*, *109*(5), 774-783. https://doi.org/10.1111/add.12470

Barratt, M. J., Lenton, S., & Allen, M. (2013). Internet content regulation, public drug websites and the growth in hidden Internet services. *Drugs: Education, Prevention and Policy, 20*(3), 195-202. https://doi.org/10.3109/09687637.2012.745828

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Becker, G.S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy, 76*(2), 169-217. https://doi.org/10.1086/259394

Chu, B., Holt, T. J., & Ahn, G. J. (2010). *Examining the creation, distribution, and function of malware on-line.* National Institute of Justice, Washington, DC.

Clarke, R.V. (1997). *Situational Crime Prevention: Successful Case Studies* (2nd ed.). Guilderland, NY: Harrow and Heston.

Clarke, R.V., & Cornish, D.B. (1985). Modeling offenders' decisions: A framework for research and policy. *Crime and Justice, 6,* 147-185. https://doi.org/10.1086/449106

Clement, J. (2019). *Global retail e-commerce market size 2014–2023*. Statista. Retrieved May 19, 2020, from https://shoptech.media/wp-content/uploads/2019/09/worldwide-retail-e-commerce-sales

Coleman, G. E. (2012). Phreakers, hackers, and trolls: The politics of transgression and spectacle. In M. Mandiberg (Eds.), *The Social Media Reader* (pp. 99-119). New York University Press.

Coleman, G. E. (2013). *Coding freedom: The ethics and aesthetics of hacking.* Princeton University Press.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management, 37*(1), 39-67. https://doi.org/10.1177/0149206310388419

Cooke, E., Jahanian, F., & McPherson, D. (2005). The zombie roundup: Understanding, detecting, and disrupting botnets. *SRUTI*, *5,* 39-44.

Copeland, C., Wallin, M., & Holt, T. J. (2020). Assessing the practices and products of Darkweb firearm vendors. *Deviant Behavior*, *41*(8), 949-968. https://doi.org/10.1080/01639625.2019.1596465

Cornish, D. B., & Clarke, R. V. (1987). Understanding crime displacement: An application of rational choice theory. *Criminology, 25*(4), 933-948. https://doi.org/10.1111/j.1745-9125.1987.tb00826.x

Cornish, D. B., & Clarke, R. V. (2003). Opportunities, precipitators and criminal decisions: A reply to Wortley's critique of situational crime prevention. *Crime Prevention Studies, 16,* 41-96.

Cox, J. (2016). Staying in the shadows: the use of bitcoin and encryption in cryptomarkets. In EMCDDA (Eds.), *Internet and Drug Markets* (pp. 41-48). Luxembourg: Publications Office of the European Union.

Cross, J.C. (2000). Passing the buck: Risk avoidance and risk management in the illegal/informal drug trade. *International Journal of Sociology and Social Policy, 20*(9/10), 68-94. https://doi.org/10.1108/01443330010789232

Cullen, F. T., Agnew, R., & Wilcox, P. (2014). *Criminological theory: Past to present, essential readings*. Oxford University Press.

Cunliffe, J., Martin, J., Décary-Hétu, D., & Aldridge, J. (2017). An island apart? Risks and prices in the Australian cryptomarket drug trade. *International Journal of Drug Policy*, *50*, 64-73. https://doi.org/10.1016/j.drugpo.2017.09.005

Décary-Hétu, D., & Dupont, B. (2013). Reputation in a dark network of online criminals. *Global Crime*, *14*(2-3), 175-196. https://doi.org/10.1080/17440572.2013.801015

Décary-Hétu, D., & Giommoni, L. (2017). Do police crackdowns disrupt drug cryptomarkets? A longitudinal analysis of the effects of Operation Onymous. *Crime, Law and Social Change*, *67*, 55-75. https://doi.org/10.1007/s10611-016-9644-4

Décary-Hétu, D., & Laferrière, D. (2015). Discrediting vendors in online criminal markets. In A. Malm & G. Bichler (Eds.), *Disrupting Criminal Networks: Network Analysis in Crime Prevention* (pp. 129-152). Boulder, Colorado: Lynne Rienner Publishers.

Décary-Hétu, D., & Leppänen, A. (2016). Criminals and signals: An assessment of criminal performance in the carding underworld. *Security Journal*, *29*, 442-460. https://doi.org/10.1057/sj.2013.39

Décary-Hétu, D., Paquet-Clouston, M., & Aldridge, J. (2016). Going international? Risk taking by cryptomarket drug vendors. *International Journal of Drug Policy*, *35*, 69-76. https://doi.org/10.1016/j.drugpo.2016.06.003

Decker, S. H., & Kohfeld, C. W. (1985). Crimes, crime rates, arrests, and arrest ratios: Implications for deterrence theory. *Criminology*, *23*(3), 437-450. https://doi.org/10.1111/j.1745-9125.1985.tb00349.x

Demant, J., Munksgaard, R., Décary-Hétu, D., & Aldridge, J. (2018). Going local on a global platform: A critical analysis of the transformative potential of cryptomarkets for organized illicit drug crime. *International Criminal Justice Review, 28*(3), 255-274. https://doi.org/10.1177/1057567718769719

Dingledine, R., Mathewson, N., & Syverson, P. (2004). *Tor: The second-generation onion router*. Naval Research Lab. Washington, DC.

Dolliver, D. S. (2015). Evaluating drug trafficking on the Tor Network: Silk Road 2, the sequel. *International Journal of Drug Policy*, *26*(11), 1113-1123. https://doi.org/10.1016/j.drugpo.2015.01.008

Dupont, B., Cote, A.M., Boutin, J.I., & Fernandez, J. (2017). Darkode: Recruitment patterns and transactional features of "the most dangerous cybercrime forum in the world." *American Behavioral Scientist, 61*(11), 1219-1243. https://doi.org/10.1177/0002764217734263

Dupont, B., Côté, A. M., Savine, C., & Décary-Hétu, D. (2016). The ecology of trust among hackers. *Global Crime*, *17*(2), 129-151. https://doi.org/10.1080/17440572.2016.1157480

Eysenbach, G., & Till, J. E. (2001). Ethical issues in qualitative research on internet communities. *British Medical Journal*, *323*(7321), 1103-1105. *https://doi.org/10.1136/bmj.323.7321.1103*

Federal Bureau of Investigation. (2012). *'Bitcoin Virtual Currency: Unique Features Present Distinct Challenges for Deterring Illicit Activity'* (Doctoral dissertation, Report from the: Intelligence Assessment, Cyber Intelligence and Criminal Intelligence Section).

Flamand, C., & Décary-Hétu, D. (2019). The open and dark web: Facilitating cybercrime and technology-enabled offences. In R. Leukfeldt & T. J. Holt (Eds.), *The Human Factor of Cybercrime* (pp. 60-80). Routledge.

Franklin, J., Perrig, A., Paxson, V., & Savage, S. (2007). An inquiry into the nature and causes of the wealth of internet miscreants. In *ACM Conference on Computer and Communications Security*, *10*, 375-388. https://doi.org/10.1145/1315245.1315292

Furnell, S. (2002). *Cybercrime: Vandalizing the information society*. London: Addison-Wesley.

Gambetta, D. (2009). *Codes of the Underworld: How Criminals Communicate.* Princeton, NJ: Princeton University Press.

Gehl, R. W. (2016). Power/freedom on the dark web: A digital ethnography of the Dark Web social network. *New Media & Society*, *18*(7), 1219-1235. https://doi.org/10.1177/1461444814554900

Gibbs, J. (1975). *Crime, Punishment, and Deterrence.* New York, NY: Elsevier.

Goncharov, M. (2012). Russian underground 101. *Trend Micro Incorporated Research Paper*. https://www.trendmicro.de/cloud-content/us/pdfs/security-intelligence/white-papers/wp-russian-underground-101.pdf

Grabosky, P. (2006). *Electronic Crime*. Upper Saddle River, New Jersey: Prentice-Hall.

Haasio, A., Harviainen, J. T., & Savolainen, R. (2020). Information needs of drug users on a local dark Web marketplace. *Information Processing & Management*, *57*(2), 102080. https://doi.org/10.1016/j.ipm.2019.102080

Hair, J. F., Black, W. C., & Babin, B. J., Anderson, R. E., & Tatham, R. L. (1998). *Multivariate data analysis.* NJ: Prentice-Hall.

Hamid, A. (1998). *Drugs in America*. Gaithersburg, MD: Aspen.

Heckman, E. (2016). *What Is Complete Separation in Binary Logistic Regression?* Minitab. Retrieved March 2021: https://blog.minitab.com/en/starting-out-with-statistical-software/what-is-complete-separation-in-binary-logistic-regression

Herley, C., & Florêncio, D. (2010). Nobody sells gold for the price of silver: Dishonesty, uncertainty and the underground economy. In T. Moore, D. Pym, & C. Ioannidis (Eds.), *Economics of Information Security and Privacy* (pp. 33-53). New York: Springer. https://doi.org/10.1007/978-1-4419-6967-5_3

Higgins, K. J. (2014). Target, Neiman Marcus data breaches tip of the iceberg. *Dark Reading.* Retrieved: May 2020: https://www.darkreading.com/attacks-breaches/target-neiman-marcus-data-breaches-tip-of-the-iceberg/d/d-id/1141162

Himanen, P. (2001). *The Hacker Ethic: A Radical Approach to the Philosophy of Business*. Random House Inc.

Holt, T. J. (2007). Subcultural evolution? Examining the influence of on-and off-line experiences on deviant subcultures. *Deviant Behavior*, *28*(2), 171-198. https://doi.org/10.1080/01639620601131065

Holt, T. J. (2010). Exploring strategies for qualitative criminological and criminal justice inquiry using on-line data. *Journal of Criminal Justice Education*, *21*(4), 466-487. https://doi.org/10.1080/10511253.2010.516565

Holt, T. J. (2012). Exploring the social organisation and structure of stolen data markets. *Global Crime*, *14*(2-3), 155-174. https://doi.org/10.1080/17440572.2013.787925

Holt, T. J. (2013). Examining the forces shaping cybercrime markets online. *Social Science Computer Review*, *31*(2), 165-177. https://doi.org/10.1177/0894439312452998

Holt, T. J. (2017). Identifying gaps in the research literature on illicit markets on-line. *Global Crime*, *18*(1), 1-10. https://doi.org/10.1080/17440572.2016.1235821

Holt, T. J. (2020). Computer hacking and the hacker subculture. In T. J. Holt & A. Bossler (Eds.), *The Palgrave Handbook of International Cybercrime and Cyberdeviance* (pp. 725-742). Springer International Publishing.

Holt, T.J., Blevins, K.R., & Kuhns, J.B. (2014). Examining diffusion and arrest practices among Johns. *Crime & Delinquency*, *60*(2), 261-283.https://doi.org/10.1177/0011128709347087

Holt, T. J., & Bossler, A. M. (2016). *Cybercrime in progress: Theory and prevention of technology-enabled offenses*. London: Routledge.

Holt, T. J., Bossler, A. M., & Seigfried-Spellar, K. C. (2018). *Cybercrime and digital forensics: An introduction*. Routledge.

Holt, T. J., Chua, Y. T., & Smirnova, O. (2013). An exploration of the factors affecting the advertised price for stolen data. In *2013 APWG eCrime Researchers Summit* (pp. 1-10). IEEE.

Holt, T. J., & Dupont, B. (2019). Exploring the factors associated with rejection from a closed cybercrime community. *International Journal of Offender Therapy and Comparative Criminology*, *63*(8), 1127-1147. https://doi.org/10.1177/0306624X18811101

Holt, T. J., & Lampke, E. (2010). Exploring stolen data markets online: products and market forces. *Criminal Justice Studies*, *23*(1), 33-50. https://doi.org/10.1080/14786011003634415

Holt, T. J., Smirnova, O., & Chua, Y. T. (2016). Exploring and estimating the revenues and profits of participants in stolen data markets. *Deviant Behavior*, *37*(4), 353-367. https://doi.org/10.1080/01639625.2015.1026766

Holt, T.J., Smirnova, O., Chua, Y.T., & Copes, H. (2015). Examining the risk reduction strategies of actors in online criminal markets. *Global Crime*, *16*(2), 81-103. https://doi.org/10.1080/17440572.2015.1013211

Holt, T.J., Smirnova, O., & Hutchings, A. (2016). Examining signals of trust in criminal markets online. *Journal of Cybersecurity*, *2*(2), 137-145. https://doi.org/10.1093/cybsec/tyw007

Holz, T., Engelberth, M., & Freiling, F. (2009). Learning more about the underground economy: A case-study of keyloggers and dropzones. In M. Backes & P. Ning (Eds.), *European Symposium on Research in Computer Security* (pp. 1-18). Berlin and Heidelberg: Springer. https://doi.org/10.1007/978-3-642-04444-1_1

Huang, S. Y., & Ban, T. (2019). A topic-based unsupervised learning approach for online underground market exploration. In *18th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/13th IEEE International Conference on Big Data Science and Engineering* (pp. 208-215).

Huang, W., & Brockman, A. (2010). Social engineering exploitations in online communications: Examining persuasions used in fraudulent e-mails. In T. J. Holt (Eds.), *Crime on-line: Causes, correlates, and context* (pp. 87–112). Raleigh, NC: Carolina Academic Press.

Hutchings, A., & Holt, T. J. (2015). A crime script analysis of the online stolen data market. *British Journal of Criminology*, *55*(3), 596-614. https://doi.org/10.1093/bjc/azu106

Hutchings, A., & Holt, T. J. (2017). The online stolen data market: disruption and intervention approaches. *Global Crime*, *18*(1), 11-30. https://doi.org/10.1080/17440572.2016.1197123

Hyslip, T. S. (2020). Cybercrime-as-a-Service Operations. In T.J. Holt & A. Bossler (Eds.), *The Palgrave Handbook of International Cybercrime and Cyberdeviance* (pp. 815-846). Palgrave Macmillan.

Hyslip, T. S., & Holt, T. J. (2019). Assessing the capacity of DRDoS-For-Hire services in cybercrime markets. *Deviant Behavior*, *40*(12), 1609-1625. https://doi.org/10.1080/01639625.2019.1616489

Ianelli, N., & Hackworth, A. (2005). Botnets as a vehicle for online crime. *CERT Coordination Center*, *1*(1), 28.

IBM (2020). Cost of a data breach report 2020. Accessed May 2021: https://www.ibm.com/security/digital-assets/cost-data-breach-report/#/

Jacobs, B.A. (1996a). Crack dealers and restrictive deterrence: Identifying narcs. *Criminology*, *34*(3), 409-431. https://doi.org/10.1111/j.1745-9125.1996.tb01213.x

Jacobs, B.A. (1996b). Crack dealers' apprehension avoidance techniques: A case of restrictive deterrence. *Justice Quarterly*, *13*(3), 359-381. https://doi.org/10.1080/07418829600093011

Jacobs, B.A. (2000). *Robbing drug dealers: Violence beyond the law*. Boston, MA: Northeastern University Press.

Jacobs, B. A., & Cherbonneau, M. (2014). Auto theft and restrictive deterrence. *Justice Quarterly*, *31*(2), 344-367. https://doi.org/10.1080/07418825.2012.660977

Jacobs, B. A., Topalli, V., & Wright, R. (2000). Managing retaliation: Drug robbery and informal sanction threats. *Criminology*, *38*(1), 171-198. https://doi.org/10.1111/j.1745-9125.2000.tb00887.x

Jacques, S., & Wright, R. (2013). How victimized drug traders mobilize police. *Journal of Contemporary Ethnography*, *42*(5), 545-575. https://doi.org/10.1177/0891241612472057

James, L. (2005). *Phishing exposed*. Elsevier.

Johnson, B.D., Dunlap, E., & Tourigny, S.C. (2000). Crack distribution and abuse in New York. *Crime Prevention Studies*, *11*, 19-58.

Johnson, B.D., & Natarajan, M. (1995). Strategies to avoid arrest: Crack sellers' response to intensified policing. *American Journal of Police*, *14*(3/4), 49-69. https://doi.org/10.1108/07358549510111947

Jordan, T., & Taylor, P. (1998). A sociology of hackers. *The Sociological Review*, *46*(4), 757-780. https://doi.org/10.1111/1467-954X.00139

Jøsang, A., Ismail, R., & Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision Support Systems*, *43*(2), 618-644. https://doi.org/10.1016/j.dss.2005.05.019

Kagan, J. (2020). *Chip Card*. Investopedia. Retrieved May 2021: https://www.investopedia.com/terms/c/chip-card.asp.

Kang, R., Brown, S., & Kiesler, S. (2013, April). Why do people seek anonymity on the internet? Informing policy and design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2657-2666). https://doi.org/10.1145/2470654.2481368

Kitchin, H. A. (2003). The Tri-Council Policy Statement and research in cyberspace: Research ethics, the Internet, and revising a 'living document'. *Journal of Academic Ethics*, *1*, 397-418. https://doi.org/10.1023/B:JAET.0000025671.83557.fa

Knowles, G.J. (1999). Deception, detection, and evasion: A trade craft analysis of Honolulu, Hawaii's street crack-cocaine traffickers. *Journal of Criminal Justice*, *27*(5), 443-455. https://doi.org/10.1016/S0047-2352(99)00015-X

Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. *Sociological Methods & Research*, *33*(2), 188-229. https://doi.org/10.1177/0049124103262065

Latour, B., Jensen, P., Venturini, T., Grauwin, S., & Boullier, D. (2012). 'The whole is always smaller than its parts'– a digital test of Gabriel Tardes' monads. *The British Journal of Sociology*, *63*(4), 590-615. https://doi.org/10.1111/j.1468-4446.2012.01428.x

Layton, R., Watters, P., & Dazeley, R. (2010, October). Automatically determining phishing campaigns using the USCAP methodology. In *2010 eCrime Researchers Summit* (pp. 1-8). IEEE. 10.1109/ecrime.2010.5706698

Lee, S., Yoon, C., Kang, H., Kim, Y., Kim, Y., Han, D., ... & Shin, S. (2019). Cybercriminal minds: An investigative study of cryptocurrency abuses in the Dark Web. *Network and Distributed Systems Security (NDSS) Symposium February 2019* (pp. 24-27). https://dx.doi.org/10.14722/ndss.2019.23055

Leukfeldt, E. R., & Holt, T. J. (2019). *The Human Factor of Cybercrime*. Routledge.

Levy, S. (1984). *Hackers: Heroes of the Computer Revolution*. Garden City, NY: Anchor Press/Doubleday.

Lewman, A. (2016). Tor and links with cryptomarkets. In EMCDDA (Eds.), *Internet and Drug Markets* (pp. 33-40). Luxembourg: Publications Office of the European Union.

Li, W., & Chen, H. (2014). Identifying top sellers in underground economy using deep learning-based sentiment analysis. In *2014 IEEE joint intelligence and security informatics conference* (pp. 64-67). IEEE. 10.1109/JISIC.2014.19

Liggett, R., Lee, J. R., Roddy, A. L., & Wallin, M. A. (2020). The dark web as a platform for crime: An exploration of illicit drug, firearm, CSAM, and cybercrime markets. In T. J. Holt & A. Bossler (Eds.), *The Palgrave Handbook of International Cybercrime and Cyberdeviance* (pp. 91-116). Palgrave Macmillan.

Lüdecke, D. (2021). sjstats: Statistical Functions for Regression Models (Version 0.18.1). *Zenodo*. https://doi.org/10.5281/zenodo.1284472

Lupton, R., Wilson, A., May, T., Warburton, H., & Turnbull, P.J. (2002). *Drug markets in deprived neighbourhoods. Home Office Research Findings No. 167.* London, England: Home Office.

Lusthaus, J. (2012). Trust in the world of cybercrime. *Global crime*, *13*(2), 71-94. https://doi.org/10.1080/17440572.2012.674183

Macdonald, M., & Frank, R. (2017). Shuffle up and deal: Use of a capture–recapture method to estimate the size of stolen data markets. *American Behavioral Scientist*, *61*(11), 1313-1340. https://doi.org/10.1177/0002764217734262

Maimon, D., & Louderback, E. R. (2019). Cyber-dependent crimes: An interdisciplinary review. *Annual Review of Criminology*, *2*, 191-216. https://doi.org/10.1146/annurev-criminol-032317-092057

Markham, A.N. (2011). Internet research. In D. Silverman (Eds.), *Qualitative Research: Issues of Theory, Method, and Practice* (pp. 111-127). Thousand Oaks, CA: Sage.

Martin, J. (2014a). *Drugs on the dark net: How cryptomarkets are transforming the global trade in illicit drugs*. New York, NY: Springer.

Martin, J. (2014b). Lost on the Silk Road: Online drug distribution and the 'cryptomarket'. *Criminology & Criminal Justice*, *14*(3), 351-367. https://doi.org/10.1177/1748895813505234

Maruna, S. (2012). Elements of successful desistance signaling. *Criminology & Public Policy*, *11*(1), 73-86. https://doi.org/10.1111/j.1745-9133.2012.00789.x

Maurer, F. K., Neudecker, T., & Florian, M. (2017). Anonymous CoinJoin transactions with arbitrary values. In *2017 IEEE Trustcom/BigDataSE/ICESS* (pp. 522-529). IEEE.

McKenzie, J. (1999). ! nt3rh4ckt! v! ty. *Style*, *33*(2), 283-298.

McSweeney, T., Turnbull, P.J., & Hough, M. (2008). *Tackling drug markets & distribution networks in the UK* (Vol. 4, No. 11). London, England: UK Drug Policy Commission.

Meikle, G. (2002). *Future active: Media activism and the Internet*. Psychology Press.

Mitnick, K. D., & Simon, W. L. (2002). *The art of deception: Controlling the human element of security*. Indianapolis, IN: Wiley Pub.

Moeller, K., Copes, H., & Hochstetler, A. (2016). Advancing restrictive deterrence: A qualitative meta-synthesis. *Journal of Criminal Justice*, *46*, 82-93. https://doi.org/10.1016/j.jcrimjus.2016.03.004

Moore, T., Clayton, R., & Anderson, R. (2009). The economics of online crime. *Journal of Economic Perspectives*, *23*(3), 3-20. https://doi.org/10.1257/jep.23.3.3

Morselli, C., Décary-Hétu, D., Paquet-Clouston, M., & Aldridge, J. (2017). Conflict management in illicit drug cryptomarkets. *International Criminal Justice Review*, *27*(4), 237-254. https://doi.org/10.1177/1057567717709498

Motoyama, M., McCoy, D., Levchenko, K., Savage, S., & Voelker, G. M. (2011). An analysis of underground forums. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference* (pp. 71-80). Berlin, Germany: ACM.

Munksgaard, R., Demant, J., & Branwen, G. (2016). A replication and methodological critique of the study “Evaluating drug trafficking on the Tor Network”. *International Journal of Drug Policy*, *35*, 92-96. https://doi.org/10.1016/j.drugpo.2016.02.027

Nagin, D. S., & Pogarsky, G. (2001). Integrating celerity, impulsivity, and extralegal sanction threats into a model of general deterrence: Theory and evidence. *Criminology*, *39*(4), 865-892. https://doi.org/10.1111/j.1745-9125.2001.tb00943.x

Nakamoto, S. (2019). *Bitcoin: A peer-to-peer electronic cash system*. Manubot. https://bitcoin.org/bitcoin.pdf

National Technology Security Coalition. (2020). *The 2020 Cyber Security Report*. Check Point Research. Retrieved August 2020: https://research.checkpoint.com/2020/the-2020-cyber-security-report/

Newman, G. R., & Clarke, R. V. (2003). *Superhighway robbery: Preventing e-commerce crime*. Cullompton: Willan Press.

Ngo, F. T., Agarwal, A., Govindu, R., & MacDonald, C. (2020). Malicious Software Threats. In T. J. Holt & A. Bossler (Eds.), *The Palgrave Handbook of International Cybercrime and Cyberdeviance* (pp. 793-813). Palgrave Macmillan.

Olivier, J., & Norberg, M. M. (2010). Positively skewed data: revisiting the box-cox power transformation. *International Journal of Psychological Research*, *3*(1), 68-95. https://doi.org/10.21500/20112084.846

Peretti, K. K. (2009). Data breaches: what the underground world of carding reveals. *Santa Clara Computer and High Technology Law Journal*, *25*(2), 375-413.

Piquero, A. R., Paternoster, R., Pogarsky, G., & Loughran, T. (2011). Elaborating the individual difference component in deterrence theory. *Annual Review of Law and Social Science*, *7*, 335-360. https://doi.org/10.1146/annurev-lawsocsci-102510-105404

Ponemon Institute. (2014). *The Aftermath of a Data Breach: Consumer Sentiment*. Retrieved May 2020: https://www.ponemon.org/news-updates/blog/security/the-aftermath-of-a-data-breach-consumer-sentiment.html

Ponemon Institute. (2019). *2019 Cost of a Data Breach Study*. IBM. Retrieved September 2020: https://www.ibm.com/security/data-breach

Pontell, H. N. (1978). Deterrence theory versus practice. *Criminology*, *16*(1), 3-22. https://doi.org/10.1111/j.1745-9125.1978.tb01394.x

Pratt, T. C., Cullen, F. T., Blevins, K. R., Daigle, L. E., & Madensen, T. D. (2017). The empirical status of deterrence theory: A meta-analysis. In F. T. Cullen, J. P. Wright, & K. R. Blevins (Eds.), *Taking stock* (pp. 367-395). Routledge.

Przepiorka, W. (2010). Diego Gambetta: Codes of the Underworld: How Criminals Communicate. *Rationality, Markets and Morals*, *1*(8), 9-11.

Resnick, P., Kuwabara, K., Zeckhauser, R., & Friedman, E. (2000). Reputation systems. *Communications of the ACM*, *43*(12), 45-48. https://doi.org/10.1145/355112.355122

Reyns, B. W. (2010). A situational crime prevention approach to cyberstalking victimization: Preventive tactics for Internet users and online place managers. *Crime Prevention and Community Safety*, *12*, 99-118. https://doi.org/10.1057/cpcs.2009.22

Rooney, K. (2019). *Online shopping overtakes a major part of retail for the first time ever*. CNBC. Retrieved May 2020: https://www.cnbc.com/2019/04/02/online-shopping-officially-overtakes-brick-and-mortar-retail-for-the-first-time-ever.html

Ruffing, T., Moreno-Sanchez, P., & Kate, A. (2014, September). Coinshuffle: Practical decentralized coin mixing for bitcoin. In *European Symposium on Research in Computer Security* (pp. 345-364). Springer, Cham.

Savage, M., & Burrows, R. (2007). The coming crisis of empirical sociology. *Sociology*, *41*(5), 885-899. https://doi.org/10.1177/0038038507080443

Schwartz, B. (2014). *Is the famous 'paradox of choice' a myth?* PBS NewsHour. https://www.pbs.org/newshour/economy/is-the-famous-paradox-of-choic

Schwartz, B. (2004). *The Paradox of Choice: Why More is Less*. New York: Ecco.

Scott, M.S., & Dedel, K. (2006). *Street prostitution. Problem Oriented Policing Guide Series* (2). Washington, DC: Office of Community Oriented Policing Services, U.S. Department of Justice.

Seals, T. (2014). 2014 so far: The year of the data breach. *Infosecurity Magazine*. Retrieved May 2020: https://www.infosecurity-magazine.com/news/2014-the-year-of-the-data-breach/

SelfKey. (2021). *All Data Breaches in 2019 - 2021 - An Alarming Timeline*. Retrieved May 2021: https://selfkey.org/data-breaches-in-2019/#

Smirnova, O., & Holt, T. J. (2017). Examining the geographic distribution of victim nations in stolen data markets. *American Behavioral Scientist, 61*(11), 1403-1426. https://doi.org/10.1177/0002764217734270

Smith, A., & Anderson, M. (2016). *Online Shopping and E-Commerce*. Pew Research Center: Internet, Science & Tech. Retrieved May 2020: https://www.pewresearch.org/internet/2016/12/19/online-shopping-and-e-commerce/

Söderberg, J. (2008). *Hacking capitalism: The Free and Open Source Software Movement*. Routledge.

Soska, K., & Christin, N. (2015). Measuring the longitudinal evolution of the online anonymous marketplace ecosystem. In *24th {USENIX} security symposium ({USENIX} security 15)* (pp. 33-48).

Statista. (2019). *Exchange rate of U.S. dollar to major currencies 2019*. https://www.statista.com/statistics/655224/conversion-rate-of-major-currencies-to-the-us-dollar/

Steinmetz, K. F. (2015). Craft (y) ness: An ethnographic study of hacking. *The British Journal of Criminology*, *55*(1), 125-145. https://doi.org/10.1093/bjc/azu061

Steinmetz, K. F. (2017). Ruminations on warning banners, deterrence, and system intrusion research. *Criminology & Public Policy, 16*(3), 725-735. https://doi.org/10.1111/1745-9133.12314

Steinmetz, K. F., Schaefer, B. P., & Green, E. L. (2017). Anything but boring: A cultural criminological exploration of boredom. *Theoretical Criminology*, *21*(3), 342-360. https://doi.org/10.1177/1362480616652686

Stiglitz, J. E. (2002). Information and the Change in the Paradigm in Economics. *American Economic Review*, *92*(3), 460-501. https://doi.org/10.1257/00028280260136363

Taylor, P. A. (1999). *Hackers: Crime in the digital sublime*. London, UK: Psychology Press.

Taylor, P. A. (2005). From hackers to hacktivists: speed bumps on the global superhighway?. *New Media & Society*, *7*(5), 625-646. https://doi.org/10.1177/1461444805056009

Thelwall, M. (2009). Introduction to webometrics: Quantitative web research for the social sciences. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, *1*(1), 1-116. https://doi.org/10.2200/S00176ED1V01Y200903ICR004

Thomas, D. (2002). *Hacker culture*. University of Minnesota Press.

Thomas, J. (2005). The moral ambiguity of social control in cyberspace: A retro-assessment of the 'golden age' of hacking. *New Media & Society*, *7*(5), 599-624. https://doi.org/10.1177/1461444805056008

Thomas, R., & Martin, J. (2006). The underground economy: Priceless. *The Usenix Magazine*, *31*(6), 7-16.

Topalli, V., Wright, R., & Fornango, R. (2002). Drug dealers, robbery and retaliation: Vulnerability, deterrence and the contagion of violence. *The British Journal of Criminology*, *42*(2), 337-351. https://doi.org/10.1093/bjc/42.2.337

Turgeman-Goldschmidt, O. (2008). Meanings that hackers assign to their being a hacker. *International Journal of Cyber Criminology*, *2*(2), 382-396.

Tzanetakis, M., Kamphausen, G., Werse, B., & von Laufenberg, R. (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. https://doi.org/10.1016/j.drugpo.2015.12.010

Van Hout, M. C., & Bingham, T. (2013a). 'Silk Road', the virtual drug marketplace: A single case study of user experiences. *International Journal of Drug Policy*, *24*(5), 385-391. https://doi.org/10.1016/j.drugpo.2013.01.005

Van Hout, M. C., & Bingham, T. (2013b). 'Surfing the Silk Road': A study of users' experiences. *International Journal of Drug Policy*, *24*(6), 524-529. https://doi.org/10.1016/j.drugpo.2013.08.011

VanNostrand, L.M., & Tewksbury, R. (1999). The motives and mechanics of operating an illegal drug enterprise. *Deviant Behavior*, *20*(1), 57-83. https://doi.org/10.1080/016396299266597

Verizon. (2020). *2020 Data Breach Investigations Report*. Verizon Business. Retrieved January 2021: https://www.verizon.com/business/en-gb/resources/reports/dbir/2021/masters-guide/

Wall, D. S. (2001). *Cybercrimes and the Internet*. New York: Routledge.

Wall, D. (2007). *Cybercrime: The transformation of crime in the information age* (Vol. 4). Cambridge: Polity.

Warnick, B. R. (2004). Technological metaphors and moral education: The hacker ethic and the computational experience. *Studies in Philosophy and Education*, *23*(4), 265-281. https://doi.org/10.1023/B:SPED.0000028400.55658.9e

Wehinger, F. (2011). The dark net: Self-regulation dynamics of illegal online markets for identities and related services. In *2011 European Intelligence and Security Informatics Conference* (pp. 209-213). IEEE.

Wilkinson, D., & Thelwall, M. (2011). Researching personal information on the public web: Methods and ethics. *Social Science Computer Review*, *29*(4), 387-401. https://doi.org/10.1177/0894439310378979

Wilson, E. (2019). Disrupting dark web supply chains to protect precious data. *Computer Fraud & Security*, *2019*(4), 6-9. https://doi.org/10.1016/S1361-3723(19)30039-9

Yip, M., Webber, C., & Shadbolt, N. (2013). Trust among cybercriminals? Carding forums, uncertainty and implications for policing. *Policing and Society*, *23*(4), 516-539. https://doi.org/10.1080/10439463.2013.780227

Zeng, G., & Zeng, E. (2019). On the relationship between multicollinearity and separation in logistic regression. *Communications in Statistics-Simulation and Computation*, 1-9. https://doi.org/10.1080/03610918.2019.1589511

Zhuge, J., Holz, T., Song, C., Guo, J., Han, X., & Zou, W. (2009). Studying malicious websites and the underground economy on the Chinese web. In M. E. Johnson (Eds.), *Managing information risk and the economics of security* (pp. 225-244). Springer, Boston, MA.