Automation and Disruption in Stolen Payment Card Markets

Timothy Peacock
Shape Security

Allan Friedman
George Washington University

Fraudulent use of credit cards represents a large component of cybercrime and data breach risk. This paper examines the card-not-present fraud chain, and identifies a critical link that might be vulnerable to anti-crime intervention. Purveyors of stolen card data rely on a testing step we term “refining”, which enables further downstream exploitation. Refining currently relies on automated abuse of web interfaces. Information asymmetries have prevented the payment card ecosystem's legitimate players from targeting this step in the harms chain. We build on previous work to identify and understand refining behavior: card-not-present merchants can be both victims and facilitators of fraud. Our contributions include a concrete, scalable intervention along with an analysis of the economics of this investment and an analysis of the present externalities of refining behavior.

“Results softened meaningfully.” This is how the then-CEO of Target described his firm’s financials after the firm’s massive data breach in December of 2013, in which over 100 million customers’ payment card information was obtained by cyber criminals.[1, 2] Breaches of confidential information continue to dominate headlines, and pundits and researchers alike have begun to document exactly how they harm the breached firm. Target, like many other organizations that have lost consumer credit card data, suffered direct and indirect losses in terms of reputation and recovery costs. This paper looks at the broader consequences of stolen data. How do breaches harm other players in the payment card ecosystem?

The dynamics of credit card fraud can appear complicated. We learn incomplete information from initial disclosures. Headlines distort and mislead. The consumer’s worms-eye view and reactions can distort the market. (One victim of the Target data breach refused to shop at Target—because she saw fraudulent charges on the breached account at another Target store, and felt they should have prevented the fraud there.[3]) Scholars have delved into many of the aspects involved in credit card fraud, from the impact of breaches[4,5] to the market for account info[6] to how criminals extract

[1] Elizabeth A. Harris, “Data Breach Hurts Profit at Target,” *New York Times*, February 26, 2014,

[2] Elizabeth A. Harris, “Faltering Target Parts Ways With Chief,” *New York Times*, May 5, 2014, http://www.nytimes.com/2014/05/06/business/target-chief-executive-resigns.html.

[3] Michael Riley, Ben Elgin, Dune Lawrence, and Carol Matlack, “Missed Alarms and 40 Million Stolen Credit Card Numbers: How Target Blew It,” *Bloomberg Businessweek*, March 13, 2014, http://www.businessweek.com/printer/articles/188935-missed-alarms-and-40-million-stolen-credit-card [3]-numbers-how-target-blew-it.

[4] Alessandro Acquisti, Allan Friedman, and Rahul Telag, “Is There a Cost to Privacy Breaches? An Event Study,” in *Proceedings of the 27th International Conference on Information Systems*, (New York: Curran Associates, Inc., 2006), 1563-1580.

[5] Arvind Malhotra and Claudia Kubowicz Malhotra, “ Evaluating Customer Information Breaches as Service Failures: An Event Study Approach,” *Journal of Service Research* 14*,* no. 1 (February 2011): 44-

value.[7] Because of this work, we understand a great deal about the shadowy world of credit card fraud. At the same time, as policy-makers and companies gear up to address this issue in the United States and around the world, we do not always understand how these pieces fit together.

This paper builds on previous work documenting the underground economy as well as previous work on market failures in security. We begin with an overview of the card fraud ecosystem and its disparate actors and incentives. We establish the market for stolen account information as a critical component of this ecosystem, and identify, in that realm, the absolute importance of the "refining" step for the efficient use of data, even in a vertically integrated carding operation. This step serves as the litmus paper that allows sellers and buyers to understand how their product will be valued and can be used in the future.

We find a market failure in protection, and suggest an investment in security technology that has the potential to rectify this asymmetry. We show that criminal behavior impacts two kinds of merchants affected by card fraud: merchants whose websites are used to test stolen credit cards and merchants whose sites are used to cash-out stolen cards. We then review different approaches to disrupting card fraud, and find that attacking this refining step has the most advantages, and the fewest disadvantages.

We observe how some merchants facilitate fraud against other merchants, and suggest a novel intervention that would allow merchants to protect both themselves and each other against fraud. By eliminating automated pre-sale testing of stolen credit cards, merchants correct a market failure, and reduce the expected value of stolen card data. Finally, we place this question in the context of the emerging political fight over responsibility and regulation brewing in Washington, D.C. We conclude that good policy will stem from identifying small technical and legal interventions that will drive efficient market-based solutions.

Overview of Card Fraud

The mechanics of credit card fraud are complex, with multiple actors, specializations, and dependencies. Rather than a single market, it makes sense to describe it as an ecosystem with niches and evolutionary dynamics. Still, there are a few key points to understand about how criminals are using payment cards illicitly, and how that hurts other actors.

From the criminals' perspective, it looks quite basic at first. The attacker needs to a) get the card data, and then b) leverage the card data to obtain something of value. The first step is the easy part. Turning a stolen credential into fungible value is where the complexity arises. Several authors have explored the different dynamics of how criminals organize and specialize to make this process as efficient as possible.[8] For fraud that will involve the use of a counterfeit card (card-present, or CP), criminals print the stolen data on magnetic stripes, and use the counterfeit card to purchase goods in real-world stores. Card-present's share of fraud is changing: a 2011 Federal Reserve report found that, for debit cards, card-not-present (CNP) fraud was the most

59, doi:10.1177/1094670510383409.

[6] Thomas J. Holt and Eric Lampke, "Exploring Stolen Data Markets Online: Products And Market Forces," *Criminal Justice Studies* 23, no. 1 (2010): 33-50, doi:10.1080/14786011003634415.

[7] Jason Franklin, Vern Paxson, Adrian Perrig, and Stefan Savage, "An Inquiry into the Nature and Causes of the Wealth of Internet Miscreants," in *Proceedings of the 14th ACM Conference on Computer and Communications Security,* (New York: ACM, 2007), 375-388, doi:10.1145/1315245.1315292.

[8] Franklin et al., "Inquiry," 375-388.

common type of fraud, and had grown rapidly over the past few years.[9] The migration to chip and PIN will likely push fraud further from CP to CNP: magnetic stripes are far easier to counterfeit than the secure chips used in Europay, MasterCard, and Visa (EMV) systems.

Card-not-present fraud is more complex. A merchant must be selected that allows the criminal to “cash-out” the value of the card. One option is to have goods delivered but this requires further layers of obfuscation to protect the criminal from identification. Alternatively, the criminal can select a cash-out merchant who deals in fungible intangible goods, for example, tickets to sporting events. Once the goods are purchased, the criminal can then resell the goods, receive payment into an account he or she controls, and subsequently use this account to transfer money beyond the reach of investigators.

Incidence of Harm

Things begin to look more complex when we try to map out the incidence of harm. There are five key parties in a payment card network: the cardholder, the issuing bank, the acquiring bank that represents the merchant, the merchant firm itself, and the card network that sets the standards.[10] (In the appendix, we review the roles and responsibilities of a payment card network in more detail.)

Under US law, the cardholder has minimal exposure to fraud risk: a maximum of $50 provided the cardholder can identify fraudulent transactions as such.[11] Most issuing banks do not choose to hold consumers liable for $50, as it causes customer churn, and the cost to acquire a customer can be well above that amount.[12]

Figure 1 illustrates the flow of information and money following a fraudulent transaction. Any fraudulent charges to cardholders are reimbursed by the issuing bank. In card present transactions, information flows as it does for card-not-present chargebacks, but the Issuer, rather than the merchant, is liable for reimbursing the cardholder. This is assuming that the merchant followed the correct procedures in accepting, processing, and storing the transaction. For CP fraud, issuing banks have direct incentives to reduce fraud rates: they are on the hook for whatever charges are rung up.

In card-not-present transactions, the issuer seeks reimbursement from the acquiring bank, who passes the costs along to the merchant, often by directly debiting their account. The merchants, thus, bear the majority of the costs of CNP fraud. They must absorb the direct cost of fraud in terms of lost product and sales, and must pay chargeback fees to the payment networks. These chargeback fees escalate, so a concentrated fraud attack will be even more painful.

Some costs are borne by the bank players: the issuing party bears the cost of reissuing the card ($5 per card, according to one industry survey[13]), and is also susceptible to a potential reputation loss if they admit a security incident that undermines customer confidence. The acquiring party is

[9] Board of Governors of the Federal Reserve System, *2011 Interchange Fee Revenue, Covered Issuer Costs, and Covered Issuer and Merchant Fraud Losses Related to Debit Card Transactions* (March 5, 2013), http://www.federalreserve.gov/paymentsystems/files/debitfees_costs_2011.pdf.

[10] Discover and American Express each are the Network, Issuing, and Acquiring organization.

[11] “Lost or Stolen Credit, ATM, and Debit Cards,” *Federal Trade Commission*, accessed May 7, 2014, http://www.consumer.ftc.gov/articles/0213-lost-or-stolen-credit-atm-and-debit-cards.

[12] Interview with former Visa fraud professional, March 31, 2014.

[13] National Association of Federal Credit Unions, “Credit Unions Pay High Price for Data Breaches,” *Economic and CU Monitor*, February 2014.

generally not liable, unless the merchant is unable to cover its chargeback liabilities and subsequently goes out of business.[14]

How do actors in the payment network fight fraud? There are two main tools. First, the issuer may reject transactions in real-time based on heuristics. Issuers can use geography, velocity, and hitting limits as signals to flag a purchase as anomalous. They have an incentive to be cautious with this tool, since false positives annoy their cardholders, and require expensive call centers to support complaints.

The card networks themselves offer value-added subscription services to score transactions on their probability of fraud, with the help of data from their much broader view of the transaction landscape. A false positive here annoys the subscribers—the merchants—by denying them a sale. Merchants can subscribe to services provided by the card networks that provide additional checks: Address Verification Service (AVS) is an option to authenticate transactions by collecting customer billing ZIP codes. Advanced versions of AVS allow merchants to selectively ask for a customer ZIP code when the bank thinks the transaction is more risky.[15] Consumers tend to encounter AVS at service stations. Because gas pump terminals do not collect signatures, these transactions are considered card-not-present transactions. Online merchants can also choose to require Card Verification Values (CVV2). These numbers are meant to never be retained by merchants, per Payment Card Industry Data Security Standard (PCI DSS), and so are theoretically not included in data breaches.

Finally, a merchant can reject a transaction as suspicious, with a similar cost of a lost sale. The former CTO of a major catalog retailer said the relatively low-tech technique of blacklisting mailing addresses was the single most effective tool against card-not-present fraud during his tenure in the late 1990s.

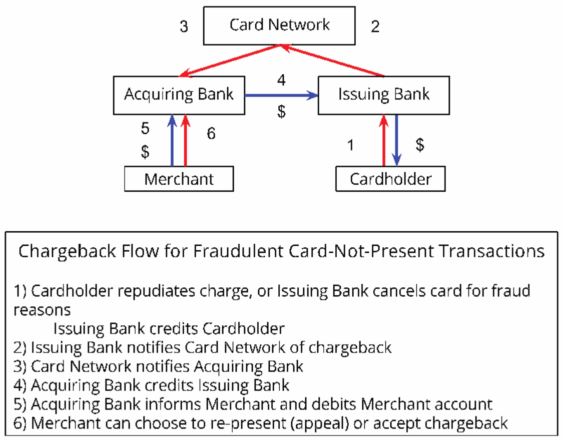
Real-time transaction rejection by merchants is useful, but does not allow data from previous transactions on that card to inform the decision—there is no capacity for state. Third party fraud scoring services are available, but also lack the longitudinal perspective. Acquiring banks have a broader view of a particular card than merchants, but only the issuing bank can view the entirety of a card's use.

In some ways, the present system aligns decision-making with the optimal data to make that decision. Only the card issuing bank can make a card cancellation decision.[16] Based on card usage data over time, a threshold can trip, and the bank will cancel the card, and "unwind" bad transactions. That is, they will deny payment and push responsibility for each fraudulent transaction onto the acquiring banks. The consumer can also aid in the detection process by reporting transactions they did not make from their own statement.

[14] Acquiring banks are also liable in cases of outright merchant fraud. We will discuss this case later.

[15] Interview with former CISO of a major global retailer, January 15, 2014.

[16] Some degree of heterogeneity has been observed in cancellation decisions. Some issuing banks cancel cards extremely proactively while others wait until fraud has happened. It is possible this is due to some issuing banks also having acquiring roles, or it could be caused by different approaches to customer relations.



Chargeback Flow for Fraudulent Card-Not-Present Transactions

1) Cardholder repudiates charge, or Issuing Bank cancels card for fraud reasons
Issuing Bank credits Cardholder
2) Issuing Bank notifies Card Network of chargeback
3) Card Network notifies Acquiring Bank
4) Acquiring Bank credits Issuing Bank
5) Acquiring Bank informs Merchant and debits Merchant account
6) Merchant can choose to re-present (appeal) or accept chargeback

Figure 1. Chargeback flow for fraudulent card-not-present transactions

The Market for Stolen Credit Cards

The marketplaces where criminals truck, barter, and exchange their wares have been descriptively and analytically studied by the security community. This work has taken many forms, beginning with some descriptive work by cybercrime scholars in the middle of the last decade. In 2006 Thomas and Martin studied a network of 35 IRC servers supporting the underground economy.[17] The next year Franklin et al. studied an IRC marketplace where they collected over 13 million messages over the course of 7 months.[18] Their marketplace was not unlike other marketplaces studied by economists: sellers advertised their wares, sometimes using automated means, buyers inquired about quality, and deals were struck. Zhuge et al. performed an extensive study of the Chinese underground web in 2008.[19]

We also know about these markets from journalists. The inestimable Brian Krebs has documented card shops, like the Rescator shop, run by a single seller or group of sellers who offer cards in a Web 2.0 e-commerce experience. The checkout system automatically checked in real-time that cards were still working.

Franklin et al. observed that selling on forums and IRC channels requires being “verified”. FBI agent Keith Mularski described the verification process to NPR reporter Zoe Chace: “In order to sell products on the site, you need to be reviewed. So if I was going to sell credit cards, what I would have to do is provide a sample of 50 cards to each reviewer. Then they would test them out and then write a review back, and say, ‘XYZ provided me 50 cards and there was a good mix of classics and platinum and business cards, and there was a 98 percent approval rating. So now I

[17] Rob Thomas and Jerry Martin, “The Underground Economy: Priceless,” ;*login* 31, no. 6 December 2006, 7-16, https://c59951.ssl.cf2.rackcdn.com/706-cymru.pdf.

[18] Franklin et al., “Inquiry,” 375-388.

[19] Zhuge Jianwei, Gu Liang, and Duan Haixin, “Investigating China’s Online Underground Economy,” (IGCC Working Paper, University of California Institute on Global Conflict and Cooperation, July 2012), http://www-igcc.ucsd.edu/assets/001/503677.pdf.

vouch for him to be a vendor on the site.'" Thus, providing quality goods is necessary to become a seller on a site.

Thomas and Martin observed channels dedicated to naming and excluding sellers who sold bad cards. Quality control is necessary to maintain status for selling on a site as well. All of the observed behaviors above point to sellers needing to perform quality control on their goods.

Pricing of Stolen Credit Cards

Reported prices for stolen credit cards have varied over time. In 2008 Symantec reported prices for credit card numbers between $0.50 and $12. More recently in late 2013 Dell SecureWorks wrote that Visa and MasterCard prices were $4 for US cards but as high as $15 for European cards. Krebs documented cards selling between $26 and $48 immediately following the Target breach that dropped to between $8 and $28 by February.[20]

What we know about prices is a lot of what we know we do not know. To begin with, the methodologies of studies like Symantec[21] and Dell SecureWorks[22] are often opaque or not disclosed at all. Their work is not repeatable or falsifiable, but it is taken at face value in multiple academic works. We also have very little visibility into prices within vertically integrated firms, or how prices change over time.

Stale Cards

Sellers of cards have a powerful incentive to sell cards as quickly as possible: the longer they hold onto a batch, the more cards within it are cancelled. Krebs documented the declining valid rate of bases (criminal terminology for batches) of cards from the Target breach.

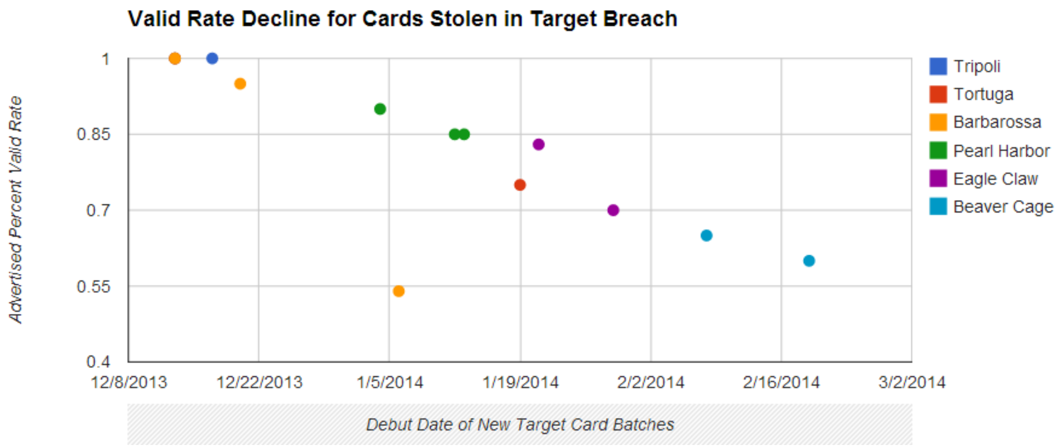


Figure 2. Valid rate decline for cards stolen in Target breach

We term this effect "going stale" and it is not limited to the Target breach: Krebs documented it

[20] Brian Krebs, "Non-US Cards Used at Target Fetch Premium," *Krebs on Security* (blog) December 22, 2013, http://krebsonsecurity.com/2013/12/non-us-cards-used-at-target-fetch-premium/.

[21] "Underground Economy Servers—Goods and Services Available for Sale," *Symantec*, accessed April 3, 2014, http://www.symantec.com/threatreport/topic.jsp?id=fraud_activity_trends&aid=underground [21]_economy_servers.

[22] "The Underground Hacking Economy is Alive and Well," *Dell SecureWorks*, accessed April 3, 2014, http://www.secureworks.com/resources/blog/the-underground-hacking-economy-is-alive-and-well/.

as well for a previous breach of Harbor Freight.[23]

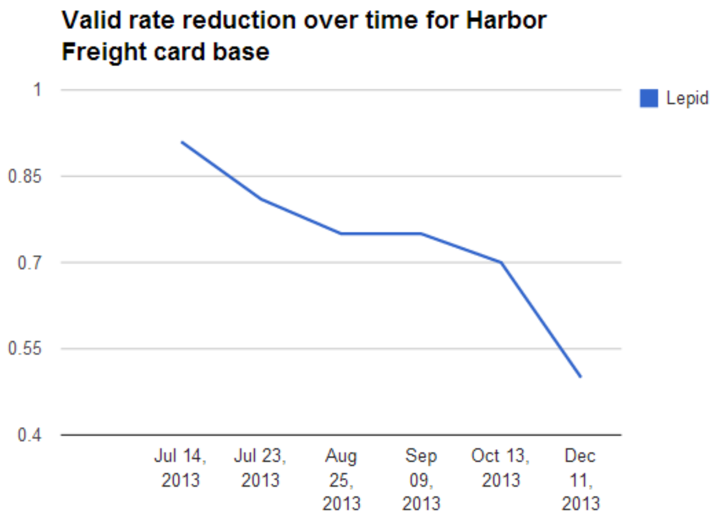


Figure 3. Valid rate reduction over time for Harbor Freight card base

Staleness is a result of several mechanisms. Issuing banks may aggressively pursue common point of purchase analysis to cut fraud before cards are cashed-out. Consumers may choose to cancel cards proactively to trade one hassle-factor for another. Staleness is a problem for cards sold inside of vertically integrated firms as well: they are subject to the same exogenous factors as card shops. Staleness affects bases in the aggregate: without testing the sellers of stolen payment card data have no insight as to which cards are still valid.

Refining—The Critical Link

We argue that pre-sale testing is a necessary step in the value extraction process though one that has had less discussion in papers on this ecosystem so far. It is a step hard to document from previous methodologies: testing of cards happens before the advertising and negotiation that can be readily observed in IRC channels.

This testing can be termed "*refining*", and we define it to be any automated process used at scale to separate working sets of payment card data. Payment card data can come from breaches of merchant point of sale systems like Target, web attacks on payment processors like Heartland Payment Systems, malware like Zeus that grabs forms as users submit to websites, phishing schemes, or real-world perfidy. This refining step allows the seller to differentiate their cards in terms of a quality signal.

Data from any of these sources can be incomplete, messy, or contain errors as simple as a flipped bit. Criminals selling cards do not necessarily know which are good: testing is necessary not only to sort out cancelled cards, but also to refine cards from key logging Trojans like Zeus. Even for card details that are captured accurately, criminals must face the exogenous staleness factors described above.

[23] Brian Krebs, "Fire Sale on Cards Stolen in Target Breach," *Krebs on Security* (blog), February 19, 2014, http://krebsonsecurity.com/2014/02/fire-sale-on-cards-stolen-in-target-breach/.

To test cards, sellers use botnets to automatically make small-value transactions on high-volume websites. By making a small purchase on a popular website, the sellers avoid tripping velocity and anomaly-based fraud detection systems. Netflix processes tens of thousands of new users each day,[24] making each sign-up nearly suspicionless to issuing banks.

Some have claimed that scammers make use of charities to test cards, including security firm Symantec.[25] The idea seems reasonable: Red Cross and other non-profits probably spend less on securing their websites and simply accept fraud as a cost of business. However, several factors work against this idea being a likely methodology. For one, charities and small businesses process significantly smaller volumes of data. Anomalous spikes in their volume will alert their teams as well as the card networks and banks. The scammers themselves seem to have some moral qualms against it as well. One self-identified former carder said in a Reddit AMA, "DO NOT DO THIS! …Donating to a charity with a stolen card will cost the charity money in chargeback fees."[26]

Criminals selling stolen cards use testing to the point that its marginal cost is equal to its marginal benefit. The marginal benefit of testing cards is an increase in the price that sellers can get for their cards.

Like other parts of the underground economy, testing is available as a service. We visited http://www.ccchecker.ru/, and registered an account. The site claims to test sets of information against AVS and CVV2 in batches starting as small as 100 or as large as 2000. Naturally, there is a discount for volume, and payment is accepted in bitcoins. Fraud services have been consumerized.

Automation is a necessary part of this process. If the attacker worked by hand testing two cards a minute, he would have to work 83 hours to test a batch of 10,000 cards, which is an order of magnitude smaller than the batch sizes observed on the Rescator shop. However, with a rented 1000 node botnet and knowledge of 10 sites that process transactions with low risk to having the cards then flagged as stolen, the seller could test 10,000 in moments. If the seller wants to test cards in real-time pre-purchase, like the Rescator shop does, automation is certainly required.

Unlike many click-farm operations like advertising fraud or Facebook "Likes" fraud, testing cards does not lend itself to being outsourced. Just as it is easier to use sweatshop labor to manufacture fake Gucci handbags than it is to use sweatshop labor to counterfeit $100 bills, we suspect a click-farm operator would have a hard time not seeing issues of theft with a human-powered testing operation.

Operational Efficiency in Disrupting Market for Stolen Credit Cards

Our proposed intervention is motivated by seeking economies of scale through operational efficiency against the harms of credit card fraud. Previous attempts to reduce credit card fraud focused on disrupting the marketplaces, which have been met with, at best, only short term

[24] Netflix added 2.33 million new users in Q4 2013. See Netflix, *Q4 13 Letter to Shareholders* (January 22, 2014), accessed April 14, 2014, http://files.shareholder.com/downloads/NFLX/3092705498x0x720306 [24]/119321bc-89c3-4306-93ac-93c02da2354f/Q4%2013%20Letter%20to%20shareholders.pdf.

[25] Yazan Gable, "Scammers Make Friends with Charities," *Security Response* (blog), *Symantec*, updated January 23, 2014, http://www.symantec.com/connect/blogs/scammers-make-friends-charities.

[26] Driverdan, "IAmA Former Credit Card Fraudster, Identity Thief, Hacker and Document Forger. AMA," *Reddit* (thread), August 28, 2013, http://www.reddit.com/r/IAmA/comments/1laau9/.

success.[27] Rather than fight an iterative battle against entrepreneurs looking to exchange goods on an Internet with unlimited dark corners, our proposed intervention interrupts the value-extraction process for all sellers of stolen cards.

Scale is critical in dealing with highly automated crime. With botnets, for example, cleaning up infected endpoints has been relatively unsuccessful. The most successful efforts against botnets have been economy-of-scale focused efforts to takedown command and control servers. One of the largest impacts on spam volume was not blacklisting of sending email addresses, but coordinated efforts to de-peer a single ISP: McColo.[28] Other efficiency gains rely on raising the cost to the attacker, or lowering the costs to the defender.

This section will consider several ways to reduce the harms of stolen credit cards. We will consider the status quo of issuing bank fraud detection, then two approaches to disrupting marketplaces with "anti-reputation" attacks. We then argue that targeting the automated services critical to large-scale crime is the optimal path.

Bank Cancellation

Issuing banks engage in fraud scoring to prevent fraudulent use of their cards. These scores are based on real-time transaction information and statistical modelling of normal card use.[29] Some vendors offer velocity-based detection, or how often a card is used.[30] Protecting against fraud by velocity or anomaly detection is akin to mopping up after blood has splattered. It is inherently reactive and uses defrauded merchants as the petri dish for detection.

The other major fraud detection and remediation technique is known as common point of purchase analysis. When an issuing institution begins to cancel cards for tripping velocity detection, or consumers report fraud, the bank can analyze the purchasing histories of the affected cards. Common point of purchase (CPP) analysis is less useful for institutions with smaller visibility into the overall ecosystem, and can potentially lead to false positives when the affected merchant is large enough.[31] CPP analysis allows banks to figure where the horse left the barn, and may inform which other horses also are missing, but like velocity measures, CPP analysis relies on some amount of fraud occurring before harm reduction can begin.

Anti-Reputation

Two attacks have been proposed against the forum and IRC marketplaces for stolen credit cards, Sybil and Slander attacks.[32, 33] These attacks are analogous to law enforcement efforts to arrest forum operators: the efforts poison marketplaces, rather than the market. As with the markets for

[27] Joseph Menn, *Fatal System Error* (New York: PublicAffairs, 2010).

[28] Dan Goodin, "Net Provider Accused of Coddling Crooks Yanked Offline," *The Register*, November 12, 2008, http://www.theregister.co.uk/2008/11/12/mccolo_goes_silent/.

[29] Banks do not tend to disclose their implementation, or even the variables involved, but it has been the subject of some academic research, such as Abhinav Srivastava, Amlan Kundu, Shamik Sural, and Arun K. Majumdar, "Credit Card Fraud Detection Using Hidden Markov Model," *IEEE Transactions on Dependable and Secure Computing* 5, no. 1 (January-March 2008): 37-48.

[30] An anonymized case-study of velocity detection: see Equifax, *Credit Card Fraud Detection Capabilities* (Atlanta: Equifax, 2012), http://www.equifax.com/pdfs/corp/0205-12_EFX-USA-2051
[30]_Suspicious_ID_Credit_Card_Case_Study.pdf.

[31] Brian Krebs, "Breach Blind Spot Puts Retailers on Defensive," *Krebs on Security* (blog), February 28, 2014, https://krebsonsecurity.com/2014/02/breach-rumor-mill-puts-retailers-on-defensive/.

[32] Thomas and Martin, "Underground," 7-16.

[33] Franklin et al., "Inquiry," 375-388.

drugs, this leads to a balloon effect: stamp out coca production in Colombia and more grows in Peru.[34, 35] When alleged Silk Road operator Ross Ulbricht was arrested, multiple Silk Road replacements sprung up.[36]

In Sybil attacks, an organization creates accounts in marketplaces for stolen credentials, gets the accounts verified, and then reneges on a number of deals. This attack is proposed to have the impact of disrupting the market by undermining buyer trust in verified sellers. Franklin et al. suggest this could be a way to "lemonize" the markets.

In Slander attacks, an actor creates many seller accounts in a marketplace. Those buyer accounts claim that sellers did not deliver goods even if the seller had. This attack is expected to undermine trust in the ability of forum admins to vet sellers, and decrease the value of a verified "nick".

As marketplace focused interventions, there are some limitations. Slander attacks could be effective against card shops, but Sybil attacks would not work against these. Sybil and Slander attacks are also not likely to be effective against vertically integrated firms. Herley and Florencio note that criminals already do these things of their own accord and have yet to break the market. Finally, none of the proposals address which actors in the payment card ecosystem would pay for this intervention. While it is an intervention that benefits everybody, it is unclear who receives the most benefit, and how those costs could be shared.

Anti-Automation

To reduce the incidence of harms to merchants from stolen credit cards, we propose that merchants deploy anti-automation technology on websites that accept credit cards. Anti-automation technology has the potential to fully prevent pre-sale refining, or raise the costs of pre-sale refining above its marginal benefits.

Longer test periods will also force criminals to sell in smaller batches, or to sell in batches with less than complete pre-sale testing. Buyers will have less confidence that each card is tested, or they will have to accept that by the time they cash-out the cards, only five of ten cards may still be working.

Impact of Anti-Automation

Implementing sufficient anti-automation for websites would protect merchants against refining behavior. As we argued in the previous section, going from an automated regime to manual testing raises the cost per card by a significant margin, and breaks the scale at which both individuals and organized rackets can operate.

This change is notable not that it dramatically increases the monetary cost of testing cards, but also by the time the seller has finished testing the batch of 10,000 cards, the results of his earliest tests are less reliable. This impact affects sellers of credit cards in all levels of the underground economy: from IRC and forums to card shops to integrated gangs. It is marketplace independent:

[34] Michelle L. Dion and Catherine Russler, "Eradication Efforts, the State, Displacement and Poverty: Explaining Coca Cultivation in Colombia during Plan Colombia," *J. Lat Amer. Stud.* 40 (2008): 399-421, doi:10.1017/S0022216X080043.

[35] Simon Romero, "Coca Production Makes a Comeback in Peru," *New York Times*, June 13, 2010, http://www.nytimes.com/2010/06/14/world/americas/14peru.html?_r=0.

[36] Andy Greenberg, "'Silk Road 2.0' Launches, Promising a Resurrected Black Market for the Dark Web," *Forbes*, November 6, 2013, http://www.forbes.com/sites/andygreenberg/2013/11/06/silk-road-2-0 [36]-launches-promising-a-resurrected-black-market-for-the-dark-web/.

defenders need not repeat their efforts each time a new marketplace springs up.

Whether anti-automation will fundamentally break the markets remains to be seen. We do expect it to have, at a minimum, an impact on the price of cards. By reducing the valid rate, and eliminating the validity of estimates of the valid rate, prices will trend down as buyers' confidence is reduced.

One possible attacker evolution would be the creation of fake merchant accounts with acquiring institutions. In this case, the cost of chargeback fees are internalized by the bank responsible for enabling the fraud in the first place. As an actor with knowledge of what legitimate merchants look like, acquiring banks are well suited to identifying and stopping this evolution.

Techniques in Anti-Automation

If anti-automation would have positive effects in theory, can it be accomplished in practice? A number of techniques have been developed since the late 1990s to limit unwanted automation of websites. These tools were not designed to prevent automated card use, but rather distinguish between legitimate, intended use of websites, and abuse of website functionality for unwanted automated behavior, such as spam. They are suited to preventing card use because the techniques used in card fraud, particularly in the refining step, rely on websites operating properly. That is, anti-automation techniques are not designed to prevent use of vulnerabilities: merchants cannot patch away an exploit when the "exploit" is accepting cards, which it must do to complete orders.

The following techniques are provided as examples of how anti-automation might be accomplished and a qualitative analysis of the pros and cons of each approach. A quantitative return-on-investment analysis is beyond the scope of this paper.

-Turing Test: CAPTCHAs have been widely deployed on the web to reduce automated interactions.[37] While many basic CAPTCHA models have been defeated through computer vision advances and distributed human labor, advances continue to more complex semantic problems, like selecting cat photos,[38] and playing games.[39]

Projects like reCAPTCHA make it easy for site operators to implement a robust, scaleable, and well-maintained anti-automation solution for free.[40] Commercial operations selling CAPTCHAs exist as well: Ticketmaster was, at time of writing, using CAPTCHAs provided by Solve Media to protect their checkout process.[41] The major downside to CAPTCHAs in e-commerce are user friction and accessibility concerns.

The W3C has been commenting on accessibility issues with CAPTCHAs for nearly a decade now, and issues around accessibility only get worse as CAPTCHAs are made "harder".[42] In

[37] Louis von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford, "CAPTCHA: Using Hard AI Problems for Security," in *Advances in Cryptology—Eurocrypt 2003*, (Heidelberg: Springer, 2003): 294-311.

[38] Jeremy Elson, John R. Douceur, Jon Howell, and Jared Saul, "Asirra: A CAPTCHA that Exploits Interest-Aligned Manual Image Categorization," in *Proceedings of the 14th ACM Conference on Computer and Communications Security,* (New York: ACM, 2007), 366-374, doi:10.1145/1315245.1315291.

[39] PlayThru, Are You a Human, accessed April 6, 2014, http://areyouahuman.com.

[40] "Google reCAPTCHA," *Google*, accessed May 7, 2014, https://www.google.com/recaptcha.

[41] Ticketmaster, accessed May 7, 2014, https://www.ticketmaster.com. Solve Media, accessed May 7, 2014, https://www.solvemedia.com.

[42] Matt May, "Inaccessibility of CAPTCHA," *W3C*, November 23, 2005, http://www.w3.org/TR/turingtest/.

addition to the impact on the visually impaired, CAPTCHAs generally annoy and inconvenience users, hurting brand reputation of organizations. If faced with challenges each time, users are likely to shop elsewhere. Organizations deploying CAPTCHAs will have to ensure that the benefits of reduced fraud costs outweigh the impact on users.

-Reputation: Reputation methods are based on historic patterns of activity from an endpoint. The central challenges are to establish a sticky and unique identity for an endpoint, and to establish whether or not an action taken by it was malicious. Common approaches to the first problem use characteristics like IP address, cookies, and fingerprinting.[43] Solving the second problem requires some amount of heuristics, and in many cases may be application specific, requiring the defender to learn what "normal" behavior looks like before protection can take place. A reputation service has the potential to be effective, but risks denying transactions through false positives when many users are routed through a Network Address Translation (NAT), as they may be at a university.

From the business model perspective, merchants could subscribe to reputation services to protect against fraud in a business model very similar to how they may currently subscribe to AVS and other card network provided fraud scoring services.[44] Having a familiar fraud scoring service to plug into their existing anti-fraud solution places the financial and administrative burden of anti-automation on a team familiar with the ecosystem.

-JavaScript Proof of Work (PoW): Web operators can require visitors to solve proof-of-work problems to make requests. This requirement increases the time that requests take to generate, thus slowing an automated attacker. Unfortunately, PoW is difficult with a heterogeneous user population and a distributed attacker.[45] Some have argued PoW is more effective when rolled together with Reputation systems.[46]

PoW work may not be the best approach, but it has the benefit of being relatively simple for a fraud team to implement. If verification is handled by a simple web server plugin, the PoW system can automatically drop requests submitted without any administrative burden on fraud or incident response teams.

-Real-Time Polymorphism (RTP): Recently polymorphic web content has been proposed as general purpose anti-automation. By rewriting web pages for each page served, web operators can prevent scripted interaction.[47] Similar to address space layout randomization (ASLR) for host defense, real-time polymorphism could break the ability of attackers to know where to interact with a web page. Just as buffer overflows may remain in an ASLR-protected world, the means to exploit them, predictable memory locations, are taken away in RTP websites by removing

[43] Peter Eckersley, "How Unique Is Your Web Browser?," in *Proceedings of the 10th International Conference on Privacy Enhancing Technologies,* eds. M. J. Atallah and N. J. Hopper, (Berlin, Heidelberg: Springer-Verlag, 2010): 1–18.

[44] Added costs for new users is a known problem with reputation systems: see Eric J. Friedman and Paul Resnick, "The Social Cost of Cheap Pseudonyms," August 11, 1999, http://presnick.people.si.umich.edu [44]/papers/identifiers/081199.pdf.

[45] For general PoW not working: see Ben Laurie and Richard Clayton, "'Proof-of-Work' Proves Not to Work Version 0.2," September 12, 2004, http://www.cl.cam.ac.uk/~rnc1/proofwork2.pdf.

[46] Debin Liu and L Jean Camp, "Proof of Work Can Work," March 23, 2006, http://weis2006.econinfosec.org/docs/50.pdf.

[47] Xinran Wang, Tadayoshi Kohno, and Bob Blakley, "Polymorphism as a Defense for Automated Attack of Websites," in *12th International Conference on Applied Cryptography and Network Security*, (forthcoming).

predictable mappings between HTML names and values. Because RTP requires no judgements about user activity, it has no potential for false positives.

This last approach differs from others in that it does not seek to detect or profile attacker behavior: the effort is to deflect automation by making it algorithmically difficult or impossible. This kind of technology could be purchased or built by hand for a particular application. Other projects, like the OWASP AppSensor have looked at web apps that are designed to defend themselves.[48] Changes to HTML to replace static references must be made across related Cascading Style Sheets and any JavaScript to avoid breaking application functionality. Making changes to web pages on the fly is difficult: it may be easier to build this kind of defense into the web applications themselves.

RTP is by far the most technically challenging of these approaches. An operation as complex as modifying entire sets of web content would most likely require cooperation across different teams within an organization. From an implementation perspective RTP may be a non-starter. From a strict “best security” point of view, however, it may be the best choice against automated attacks.

Finally, merchants could shift to using third-party payment processing services. Services, like Stripe,[49] provide JavaScript based forms to collect, process, and manage credit and debit card payments. By serving as payment processor to thousands of small websites, such services can build in anti-automation technology and use it with all of their customers.

Choosing an anti-automation solution that is right for a firm will be function of many factors, from organizational design (how much authority does the fraud team have to install new technology) to budget (free reCAPTCHA use vs. yearly reputation service subscription vs. real-time polymorphism hardware) to concerns over user friction and false positive rates (high impact of CAPTCHAs to zero false positives of RTP).

Externalities

Underlying the field of security economics is the notion that we can arrive at better public policy by re-aligning market forces to correct externalities. This discussion of credit card fraud helps us identify two glaring externalities. The first, more evident, is the misalignment between the issuing banks’ cancellation decisions and the Internet, mail order, or telephone order (IMOTO) merchant’s exposure to fraudulent purchases. More subtly, the above discussion of carder “refining” helps us understand a split in capacity and responsibility between different types of merchants in the fraud ecosystem. This section details these externalities, and proposes a solution to address them.

For the relationship between the issuing bank and the merchant, recall from the previous section, “Overview of Card Fraud” that an IMOTO merchant largely bears the responsibility for card-not-present fraud. In most cases, the merchant is left holding the bag for fraud that has been identified after-the-fact as fraudulent. This is seen as the cost of doing business, but it is also something the merchant has little control over. It is not a completely exploitative relationship, of course. The market functions because the merchant can make some decisions about what kind of transaction to accept or reject, and when to demand more information that might lessen the likelihood of fraud.

[48] OWASP AppSensor Project, *OWASP*, accessed April 6, 2014, https://www.owasp.org/index.php
[48]/OWASP_AppSensor_Project.
[49] Stripe, accessed May 7, 2014, https://www.stripe.com.

We can compare this to the dynamics of fraudulent card-present transactions. In this case, the issuers bear the costs of fraud following a properly authorized transaction, paying the aquiring bank and merchant out of pocket. This also presents an externality if the merchant does not internalize the costs of fraud, but the banks are not helpless. Since they bear the brunt of the responsibility, issuing banks monitor merchan behavior, charging penalties to those responsible for too many chargebacks. The merchant is in a position to detect egregious fraud behavior, and try to mitigate it. For example, a store can also ask for photo ID, a mild form of biometric authentication, that either limits fraud options or raises the cost of bulk value extraction. More generally, merchants must optimize between allowing fraudulent transactions through (a type II error, or false negative) and preventing a legitimate transaction (a type I error, or a lost sale and an irked customer). In the status quo, the merchant must make this decision alone.

Do payment card accepting merchants have the support they need to best make this decision? Some basic solutions have been proposed, but are seen as failures. Chip and PIN protection makes card-present fraud hard, although not impossible (see the next section), and additional data demands from the merchant or payment companies have been shown to either be expensive, onerous to customers, or ineffective.[50] Since they do not have any direct determination to cross correlate multiple purchases on a single card to identify either a "refining" purchase or a value extraction purpose, they must look at the transactions themselves.

There are some efficiencies for placing the transaction rejection decision in the hands of the merchant. They have the ability to determine their preference over the costs of false positives and false negatives, and can set their risk level accordingly. But they still suffer from lack of data, and need more defense-in-depth. Anti-automation tools allow the merchants to mitigate some of the risks from automated purchase fraud, and avoid the costs of chargebacks, which are around $25 per transaction. Recall that chargebacks result in fees, which can grow high over time if a merchant is a frequent victim. Reducing chargeback fees will help all merchants.

It is important to remember that there are really two types of merchants in the refined-card model: the *refining merchant* and the *cash-out merchant*. The cash-out merchants bear a much greater absolute risk, since they have the most value at stake. IMOTO cash-out merchants and the issuing banks held responsible for card-present fraud need some mechanism to encourage refining merchants to adopt anti-automation technology. There is an information asymmetry, since the refining merchants have no direct visibility that their own victimization is imposing an even higher cost on others. The cash-out victim also has no view of the upstream role played by the refining merchant.

As such, it is not hard to see that adoption of anti-automation technology by frequently-targeted refining merchants could help address this negative externality. The refining merchant gains some benefits, but may not have the necessary incentives to invest to a point that could optimally protect the cash-out merchant. The latter has no direct mechanism to identify the former for an efficient side payment. How can we tinker with the market structure to promote efficient investment?

Regulatory Mandate

If we deem this issue to be of great enough importance, we might imagine government regulation

[50] Steven J. Murdoch and Ross Anderson, "Verified by Visa and MasterCard SecureCode: Or, How Not to Design Authentication," in *Financial Cryptography and Data Security* 6052, ed. Radu Sion (Berlin Heidelberg: Springer, 2010), 336-342.

as a solution. After all, financial regulators intervene in a host of matters about the roles and responsibilities in payment cards. If there was concrete evidence that this issue was large enough to merit regulatory attention, how might the government intervene? A regulator could require organizations who process a certain number of transactions that are deemed fraudulent to implement effective protections such as those described above, taking into account potential efficacy. The FTC's Red Flags model might be an appropriate model here, since credit card fraud is one of the warning signals of identity fraud these rules address. This also has the benefit of allowing the red flag to correlate with direct losses to the firm in question from chargebacks and compliance costs. On the other hand, this approach does not directly address the question of information asymmetry, since this approach only addresses merchant-detected fraud, not those that occur upstream. A stronger regulation might be more effective at pre-emptive detection, but would require targeting multiple parties and forcing cooperation into what has often been a competitive, zero-sum relationship. We should note that regulation is probably unlikely, for several reasons. This issue has not emerged as demanding Washington's attention, posing neither a large enough problem nor one immediately affecting consumers, two factors that drive regulatory attention. In general, any regulation that attempts to impose a technology-specific solution in a dynamic environment has poor odds of succeeding, particularly when the adversary has demonstrated the ability to adapt at a pace that far exceeds regulatory response.

Liability

If refining merchants bear costs that are less than the fees and hassles of their bulk, low level fraud, then perhaps the cash-out firms that bear higher costs later on could force accountability through liability. If the existing statutes cannot support judicial expansion of tort in this case, the government may have to step in and explicitly impose this legal relationship. Determining the exact nature of liability, and the standards necessary to avoid such legal exposure has proven extremely difficult in other areas of cybersecurity policy, such as individual responsibility for unsecured machines, or accountability for buggy software. If we can surmount this issue through appropriate standards of due care, how might liability promote information sharing? Beyond data from their own payment systems, merchants might need information about patterns of fraud that impact others. Which accounts have been involved in breaches? What other suspicious activity has been observed? This information is in the hands of the card networks and issuing banks. If merchants have a financial interest in this view to protect themselves, they might be willing to pay for it. This could thus create a market for credit card fraud behavior data. The networks and banks would have to collect financial information, and package this data to be useful to merchants, priced low enough to be attractive enough to be easily integrated into merchant systems. This assumes that the merchants themselves have a strong enough financial incentive to have a sophisticated fraud detection system. This assumption cuts both ways, however. While sophisticated, well-endowed merchants can implement such a system, many merchants that depend on credit card sales use pre-baked systems built for scale rather than merchant-specific demands, or lack the resources to implement such a data-intensive fraud system.

Cross Subsidies

How can smaller refining merchants help mitigate this risk without being priced out of the payment card industry? We have three parties: 1) acquiring banks with vertical and horizontal information on card use behavior that might be used to detect fraud, 2) cash-out merchants who are greatly harmed by the larger, targeted fraud, and 3) refining merchants who bear some costs of fraud, and high costs of detection. Can the cash-out merchants align to subsidize the process of information sharing and technical investment? If cash-out merchants can both understand their victimization, as well as the dynamic nature of their internal defenses, they may have an incentive. What is left is a collective action problem—how can they work together. The payment card networks have mastered the art of cooperation under the PCI model, but the merchants may be a diffuse group. The networks and banks might have the necessary data to identify a shared body, and use their role as an intermediary to help address some of the coordination costs. Why would they do this? In the final section, we outline a political dimension that highlights the value of market players demonstrating their value as good faith actors in a competitive space.

The Political Landscape

Crisis and media attention often drive public policy. This can lead to action, even when the best path forward has not been determined. This is true in government, but also true in the boardroom and even in technical standards organizations. The attention paid in the United States to the recent breaches of Target and Neiman Marcus pose such a risk. Both the House of Representatives and the Senate held hearings in February 2014, and many solutions were discussed. Was it time for the United States to finally adopt the chip and PIN model? Do we need to mandate stronger defenses for those holding valuable consumer data? Would shifting data breach legislation from a state issue to a national issue change anything? Even those arguing against any increase of the role of government seemed to be hinting at broader policy development: Rep. Marsha Blackburn, co-chair of the House Privacy Working Group, called for taking "the rules on the books for the physical space and apply them to the virtual space to encourage commerce."

The issue is critical for public policy, because a lot of money is at stake. Credit cards are big business. The fees charged just for using them at a point of sale total over $30 billion annually in the United States alone.[51] The question of how to deal with consumer fraud is more than a political dimension to a cybercrime problem. More than the cost of fraud, the debate is about the cost of solutions. Who will pay? Who will bear legal liability? Any deviation from the status quo or chance to change the existing equation will be engaged eagerly, and it is critical to have a clear understanding of the costs and benefits. Not only are many of the solutions discussed following 2013's high profile breaches expensive, many of them will not work.

One of the most popularly discussed solutions to the theft and misuse of credit card data is the implementation of smart cards and readers, the "chip and PIN" option. The cryptographically protected chip embedded in a credit card makes counterfeiting cards from stolen data much more expensive, and an ideally-implemented sales system would not involve any valuable stored data that could be stolen and misused. Chip and PIN has been called for frequently in the wake of the Target breach.

Unfortunately, chip and PIN is far from a perfect solution. First, it only protects transactions at a compliant point of sale. When this technology was implemented in the United Kingdom, fraud simply shifted to card-not-present venues and transactions outside the area of implementation. Fraud ultimately grew above pre-chip and PIN levels. It is also far from invincible. Researchers have demonstrated that an attacker can use a stolen card without a PIN,[52] and skim details with a corrupted reader.[53] Attempts to introduce technology to support chipped transactions at home

[51] Jennifer Bjorhus, "Judge Approves $7.25 billion Swipe Fee Settlement, But Battle Continues," *StarTribune*, December 16, 2013, http://www.startribune.com/business/236053321.html.

[52] Steven J. Murdoch, Saar Drimer, Ross Anderson, and Mike Bond, "Chip and PIN is Broken," in *2010 IEEE Symposium on Security and Privacy*, (IEEE, 2010), 433-446, doi:10.1109/SP.2010.33.

[53] Andrea Barisani, Daniele Bianco, Adam Laurie, and Zac Franken, "Chip and PIN is Definitely Broken: Credit Card Skimming and PIN Harvesting in an EMV World," (Inverse Path presentation at CanSecWest 2011), http://cansecwest.com/csw11/Chip%20&%20Pin%20-%20Barisani%20&%20Bianco.pdf.

have also been less than successful.[54]

Other policy options to address the risk of breaches and fraud may not be any more successful. One industry group representing financial institutions called for new legislation requiring "any entity responsible for the storage of consumer data to meet standards similar to those imposed on financial institutions under the Gramm-Leach-Bliley Act."[55] While better data security is a laudable goal, national data protection standards are an enormous step, given the absence of, say, similar standards to protect the operation of critical infrastructure information systems.

Since this call came from an industry group, it highlights the fact that the battle lines are drawn. If we look at information security as a zero-sum game, where responsibility and liability will be pushed onto one party or another, then this political debate over security will become a bloody battle. The retailers have geared up, taking out public ad campaigns in the airports and subways of Washington, D.C. It is not a question of political values, either. In a regulatory debate over debit cards, both sides spent millions on lobbyists, and had strong partisans on both sides of the aisle.

From a public policy perspective, how can we move towards an important and popular goal—better security and less fear of fraud—without the rancor, expense, and misleading rhetoric that accompanies acrimonious battles? The consumer protection policies, for example, have pushed the investment in fraud detection. Cambridge University's Ross Anderson has been vocal in highlighting how misplaced liability leads to underinvestment in security. This paper has argued that small interventions can drive functional market-based solutions.

Limitations and Future Work

This paper argues that refining is a critical step in the value-extraction process of credit card fraud, and that anti-automation technology can be a key tool against this, mitigating fraud overall as long as the necessary incentives are in place to promote its deployment. We acknowledge that there are two broad counters against this argument. First, the technology and economic mechanisms to promote its deployment may not actually work to stop the refining step. Alternatively, one might argue that even a successful elimination of the refining step would not have a tangible impact on the sale and use of stolen card credentials. These arguments serve to highlight the complexity and interactive dependencies of the cyber fraud space, and motivate more research questions, with the potential for data to further clarify these issues.

Why would this technology not halt the practice of refining? If they could not do it automatically, carders may switch to a manual system, exploiting cheap labor around the world in 'click farms' such as those seen countering CAPTCHAs and other security mechanisms. This is a much worse option for carders, and not just because of the increased cost. Manual refining could increase the costs of the refining step by several orders of magnitude, but it still could be cost-effective to spend that money to deliver a better product. Human power lacks the timeliness of an automatic check, meaning that vendors and buyers could not do this in real time to verify the quality of goods at the marketplace, a hurdle to trust. Moreover, the manpower itself may not be as

[54] Arjan Blom, Gerhard de Koning Gans, Erik Poll, Joeri de Ruiter, and Roel Verdult, "Designed to Fail: A USB-Connected Reader for Online Banking," in *Secure IT Systems* 7617, eds. Audun JÃžsang and Bengt Carlsson (Berlin Heidelberg: Springer, 2012), 1-16, doi:10.1007/978-3-642-34210-3_1.

[55] B. Dan Berger, "Are Retailers Doing Enough to Protect Consumers from Data Breaches?," *The Blog* (blog) *Huffington Post*, December 26, 2013, http://www.huffingtonpost.com/b-dan-berger/are-retailers [55]-doing-enoug_b_4505023.html.

trustworthy. There’s a difference between trusting factory workers to make knock-off Gucci handbags, versus counterfeiting $100 bills, or anything with significant value. Similarly, click fraud requires less trust than sending one's hard-won credit card numbers to an unknown network of anonymous, poorly paid workers. Empirically, further work on the actual mechanism of cyber sweatshops could help decide this question, as well as a better model of the precise value of timely checking in the marketplace.

Alternatively, the attackers could simply adapt by going after different merchants, out of the millions that process credit cards online. Yet attackers can’t target just anyone. There is some cost in scripting an attack against a new interface, and they really want to hide refining tests with merchants who process a large number of transactions. We have begun discussions with several entities that might offer insight into the distribution of transactions of different sizes, which could be used to build a threshold model of how to hide in such traffic.

A truly adaptive attacker might find some way to defeat the anti-automation tools technically. The question is which security technologies to prevent unwanted automation allow attackers to evolve most quickly, and which ones best allow for defender evolution? In the past, defenders have relied on finding "hard problems", as Luis Von Ahn titled the seminal computer science paper on CAPTCHAs. Unfortunately, computers have gotten faster and attackers have gotten smarter. Defense techniques that rely on hard problems have naturally limited lifetimes. Attackers can respond by working out what signals are being measured, or how the signals are being measured. For example, detection based on rate limits works until the attackers stay under the limit, or until the limit gets so low as to cause false positives. Attacker evolution is one reason that investing in real-time polymorphic (RTP) techniques may be the correct choice for defenders. RTP techniques take away the ability of an automated attacker to interact with a website at all: without the hooks of known HTML parameters, attackers have nowhere to enter their credit card numbers. Furthermore, a service provider who implementspolymorphism-as-a-service could centralize R&D costs to stay ahead of any attacker evolution. Additionally, an anti-automation solution could work in a "pass mode" that allows transactions to go through while flagging those computers as fraudulent actors. Combined with robust computer fingerprinting, this could allow a business model where anti-automation is provided for free. CNP merchants could subscribe for endpoint reputation and issuing banks could subscribe to know which cards to cancel.”

The techniques above may be successful at denying fraudsters a more efficient market by a refined card product, but still have minimal impact on the carding ecosystem itself. An unrefined card is still worth something but, as we describe above, anecdotal evidence suggests that it is worth much, much less, up to an order of magnitude or more. This will reduce supply but, we must acknowledge, offer temptations for new players to potentially enter the market on the demand side. At first, these will not be as sophisticated, but they may improve. Similarly, it will further reinforce an existing trend for vertical integration, but those organizations also derive value from a pre-cashout refining step, so their expected value will also decline.

Future work could build out the model from the carding side further down the supply chain. What is the cost of a rejected attempt to cash out a cancelled card? If the defenders at high-value targets can identify and track attacker identifying information, such as IP or machine fingerprinting, it can make value extraction much harder. Data-wise, we have had several discussions attempting to understand the relationship between card cancellation and fraudulent use. Of blocks that are stolen, what fraction were (fraudulently) used when still working and what fraction had been cancelled?

There is a final risk of introducing a new perverse incentive. If refining and cash-out merchants succeed in working together, and genuinely reduce the expected gains of CNP fraud, those who trade in stolen accounts may seek other venues to extract value. This may increase the attractiveness of card-present fraud, via counterfeit cards. As we note above, the acquiring banks bear substantial liability from counterfeit carding operations. If they anticipate a major shift in this direction, then the acquiring banks may be less willing to cooperate with online merchants for fear of bringing the problem into their own house. Fortunately, chip-and-pin defense will mitigate the long-term threat of counterfeit cards, so the economic calculation should stay focused on the CNP issue.

Conclusions

Security in the financial sector is a risk calculation, but we need to align incentives properly. We have to acknowledge that security is "part of a balanced breakfast" with no single solution. But steps forward should reflect the dangers of card-not-present fraud and the growing mobile and online payment markets.

Payment card fraud is a large and growing problem. Consumer protection laws have shielded the average user and preserved a critical commercial tool, but the costs fall on other actors. For card-present fraud, the issuing banks bear this cost, and have some means of controlling these costs through contracts and rules governing their relationships with merchants. For card-not-present fraud, the brunt of the pain from fraud falls on the online merchants. In both cases, the banks and merchants have a strong market incentive to reduce successful fraud attempts based on stolen card data. Anything that credibly and affordably reduces the value of this card data would thus be of interest to both card issuers and, perhaps even more importantly, online merchants.

The focus advocated in this paper is to target a small aspect of the carding ecosystem, but one that will cripple most of the criminal models that have been explored in previous literature. We presented several approaches that existing payment card industry actors might consider for disrupting the market for stolen credit cards. Of these, anti-automation technology appears to offer the greatest chance of success. At the moment, however,there are insufficient incentives for investment in this technology. We present several potential mechanisms to realign incentives for socially efficient investment, and analyze them in a political as well as economic context. Further empirical work is needed to more precisely characterize the impact and damages of the carder refining before any regulatory action should be considered.

Following the tenets of security economics, we advocate investments which make it harder to extract value, while empowering the defenders. Moreover, empowering merchants through novel collaborative relationships to prevent automated card usage will create local benefits by reducing their own fraud costs, driving the diffusion curve for early adopters. Criminals will continue to adapt and find new tactics, but anti-automation models will continue to be useful as both attackers and defenders evolve.

Appendix: Payment Card Processing

We included this appendix to help readers understand the different actors involved, the two kinds of payment card transactions, how information flows for a transaction, and who ends up holding the bag for a fraudulent transaction. Fee structures for ordinary transactions, rewards programs, branding and marketing etc. are not germane to this paper, and so are not discussed.

Payment Card Actors

The simplest model of the payment card ecosystem has five actors: Cardholders, Issuing Banks, Card Networks, Merchants, and Acquiring Banks.

-Cardholders: Cardholders are consumers to whom credit is issued. Vetted by the Issuing Bank to be creditworthy. Cardholders make purchases using their cards and settle their debts periodically with their Issuing Bank. Cardholders initiate transaction authorization by presenting their cards to merchants or by providing their card information (card number, expiration date, CVV2). Responsible for noticing fraudulent charges on statements. Not liable for fraudulent charges under most consumer protection laws, subject to certain time limits and caveats.

-Issuing Banks: Issuing Banks are institutions that issue credit to consumers. Responsible for vetting their customers, and are liable for debts that consumers are unable to pay. Issuing Banks confirm in real-time that cardholders have credit remaining, and provide authorization for transactions. Settle with consumers at end of billing cycles. Responsible for detecting card fraud, cancelling those cards, and rolling back fraudulent transactions. One of the early and popular fraud detection technologies was Fair Isaac Falcon. Issuing Banks are not liable for card-not present-fraud.

-Card Networks: Card Networks, like Visa and MasterCard, are responsible for providing the technology that interconnects Merchants, Acquiring Banks, and Issuing Banks. There are some situations where Card Networks also fulfill the role of Acquiring Bank and Payment Processor, but in most cases Card Networks are facilitators of transactions. The information that flows from a Merchant's Acquiring Bank to an Issuing Bank includes the Merchant ID, Merchant Category, Merchant ZIP code, Acquiring Network, and Transaction amount. Issuing Banks can also subscribe to fraud scoring services provided by the Card Network. All of this information can be used in real-time to make an authorization decision, or to flag the transaction as potentially fraudulent and escalate to support teams.

-Merchants: Merchants are businesses who accept credit cards as payment for goods or services. They maintain the technology to process credit cards, whether that is swipe machines or websites, and keep a merchant account with an Acquiring Bank. Merchants are liable for fraudulent purchases, and pay penalties for transactions that Cardholders or Issuing Banks reverse.

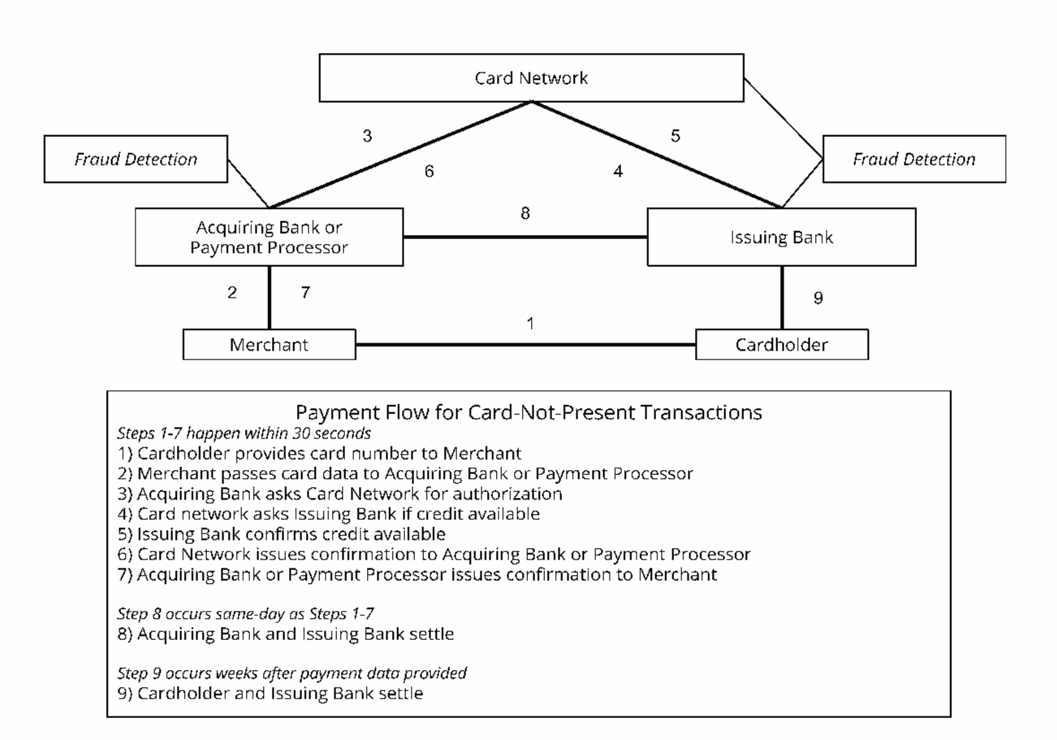
-Acquiring Banks: Acquiring Banks work with Merchants to settle transactions debits and credits. Sometimes an intermediary known as a Payment Processor sits between Merchants and Acquiring Banks to provide protection against the risk that the Merchant defaults, or has so many chargebacks the Card Network closes the account. There is the interesting case that when the Acquiring Bank is the same as the Issuing Bank, chargeback resolution occurs in-house. The non-adversarial nature of this situation leads to lower chargeback costs for the bank and the merchant. Merchants therefore have an incentive to work with Acquiring Banks who are also large Issuing Banks.

Transactions

There are two types of payment card transactions: *card-present* and *card-not-present.* Card-present (CP) transactions can be conducted by electronically reading the magnetic stripe, by manually entering the card numbers, or by physically taking an imprint of the card. The transaction is finalized by collecting the cardholder's signature either on the receipt or on an electronic terminal.

Card-not-present (CNP) transactions are not new to the Internet, having been a fixture of catalog and phone ordering for decades. In these transactions, cardholders share their payment information with the merchant without handing the merchant the card or providing a signature. The remainder of this appendix will focus on CNP transactions

We present below an archetypical CNP transaction flow. This model is somewhat simplified, and shows only one possible workflow for a payment card transaction; there are a number of other transaction processing schemes including pre-authorization, or delayed settlement. It also



Payment Flow for Card-Not-Present Transactions

Steps 1-7 happen within 30 seconds

1) Cardholder provides card number to Merchant
2) Merchant passes card data to Acquiring Bank or Payment Processor
3) Acquiring Bank asks Card Network for authorization
4) Card network asks Issuing Bank if credit available
5) Issuing Bank confirms credit available
6) Card Network issues confirmation to Acquiring Bank or Payment Processor
7) Acquiring Bank or Payment Processor issues confirmation to Merchant

Step 8 occurs same-day as Steps 1-7

8) Acquiring Bank and Issuing Bank settle

Step 9 occurs weeks after payment data provided

9) Cardholder and Issuing Bank settle

Figure 4. Payment flow for card-not-present transactions

Fraud Liability

Given that consumers cannot be held liable in the United States for fraudulent purchases made with their cards, some party in the remaining set of Issuing Banks, Acquiring Banks, and Merchant Banks ends up liable for the loss.

A chargeback is the mechanism by which an Issuing Bank reverses a transaction. Figure 1 shows a diagram of the typical workflow. The Issuing Bank tells the Acquiring Bank it is reversing a particular transaction. The Acquiring Bank debits the amount of money from the Merchant Account and credits the Issuing Bank and charges the Merchant a fee.

For the Merchant, this means that a fraudulent transaction not only has the cost of the goods or services provided to the fraudulent party, they also have to pay the chargeback fee. These fees vary depending on the processor, $15 for recent entrant Stripe, $20 for Internet stalwart PayPal, and escalating depending on the fraction of transactions reversed for others. Too high a percent of chargebacks and Merchants can see their Merchant Accounts closed. The defrauded Merchant is also out the original transaction fees.

Naturally Merchants take some steps to detect and prevent fraud. Card-Not-Present Merchants who deliver physical goods can maintain blacklists of addresses previously shipped to for fraud. One former CTO of a fashion catalog retailer said this was the single most effective measure in his fraud prevention arsenal. A decidedly low-tech solution, blacklisting addresses relied on the difficulty of scaling physical addresses.

Card Cancellation

Issuing Banks can cancel payment cards that are stolen. The cancellation decision rests with the Issuing Bank or the Cardholder. Issuing Banks create fraud scores on a number of factors, some of which are kept as trade secrets, others of which are obvious. Factors like velocity of purchasing, size of purchases, geography of purchases come into play as do others. Issuing Institutions can also subscribe to Card Network fraud scoring services.

Threat intelligence services purport to provide useful credit card scoring to Issuing Banks by monitoring the card fraud forums and IRC channels. This approach makes sense in some cases: Franklin et al. found over 100,000 unique credit card numbers in the IRC logs they examined. However, the majority of cards available for sale cannot be posted publicly: once posted in its entirety the seller cannot demand compensation for the card. These services are therefore likely identifying only cards that have already been drained, and therefore identifying the card to the Issuing Bank provides little additional value for the Bank.

The factor regarding card cancellation that is most relevant for criminals is whether the Issuing Banks are able and choose to correlate card cancellation decisions. This factor varies from bank to bank. Some banks may choose to be extremely proactive in cancelling and reissuing cards, while other banks may wait for a specific card to be used for fraud before cancelling it.

Ultimately Issuing Banks are able to recover the costs of fraud from Acquiring Banks and Merchants. Merchants are on the hook for fraud, and chargeback fees cover the costs of reissuing cards and cleaning up fraud.