This Is Your Data on Drugs:

Lessons Computer Security Can Learn From The Drug War

David Molnar
Microsoft Research
dmolnar@microsoft.com

Serge Egelman
Brown University
egelman@cs.brown.edu

Nicolas Christin
Carnegie Mellon University
nicolasc@andrew.cmu.edu

ABSTRACT

Researchers have recently begun to study the economics of the markets for *illicit digital goods* to better understand how to invest resources in the most effective mitigations. This line of work in security economics can greatly benefit from data gathering methodologies used for the study of another underground economy, which has been analyzed for the better part of a century: the illicit drug trade. We describe "promises" and "puzzles" in the use of observational data for computer security research, that have been encountered previously in drug policy research, and highlight possible lessons we can learn from this different field. We then outline potential opportunities for security research to avoid pitfalls in data collection that drug policy studies have uncovered. Finally, we argue that failure to tackle problems with observational data runs the risk of creating incorrect "mythical numbers" that can have lasting effects on public policy surrounding computer security.

1. INTRODUCTION

The last three years have seen a growing amount of security research focus on observed market prices for compromised credit cards, bank account logins, machines ("bots"), or other *illicit digital goods.* This research parallels a decades old trend in using observed market prices to draw conclusions about *illicit drugs*, such as marijuana, cocaine, or heroin. Both areas of research concern criminal behavior, where the inner workings of individuals or firms in a market are difficult to observe directly. The promise of price data is that, together with our understanding of how markets behave, prices can tell us important information about criminal behavior, even though we cannot see this behavior first hand.

Drug policy uses price information explicitly as a tool to measure the effectiveness of interventions in the market. Assuming that demand remains constant over short periods, economic theory tells us that a sharp increase in prices after an intervention in the market, such as seizing a shipment of drugs, means that intervention has successfully reduced the supply of goods. This is why the U.S. Drug Enforcement Agency issues press releases when the street price of methamphetamine spikes, pointing to such a spike as evidence that the agency's efforts to disrupt drug distribution are succeeding [10].

Price increases may also be a desirable goal by themselves for drug policy, if they lead to reduced demand for illicit drugs. Reduced demand is not only important for its own sake, but also because it may in turn lead to reducing externalities associated with illicit drug use, such as emergency room visits or theft to fund a drug habit. The key assumption underlying this goal is that the demand for drugs is elastic with prices—as prices increase, demand decreases. Given the addictive nature of drugs, the validity of this assumption is not obvious, so price data is also used by researchers to test the elasticity hypothesis [5].

Ideally, we could transfer approaches used in research on illicit drugs directly to evaluate proposed interventions in the market for illicit *digital* goods. Here "intervention" is broadly construed: an intervention could be anything from a new security technology to the arrest of a particularly vexing criminal. Measuring the effectiveness of an intervention strategy is important because it enables optimal allocation of a fixed security budget across multiple possible interventions.

From a security research perspective, measuring price data also gives us a new way to evaluate proposals. Instead of reasoning in a vacuum about the adversary's capabilities, we can perform a test deployment of the proposal and observe the effect on prices for the appropriate illicit goods. Watching how prices change over time after the deployment may also give us insight into how quickly adversaries *adapt* to new proposals, which is currently poorly understood.

Hence, observing prices offers a way out of the classic problem in evaluation of security research: we can not know the capabilities of the adversary in advance, nor can we accurately measure the rate at which the adversary adapts to new security measures. Because of our poor understanding of adversary capabilities, security researchers *must* be conservative in evaluation of new designs. This in turn biases our field towards proposals that are more difficult to implement, less usable, or incur more implementation cost than would be the case if we had a better understanding of the environment in which they are to be deployed. In Section 2 we describe how recent inquiries in security economics can defines a new paradigm for security research in terms of "promises" from using observational data. We also describe "puzzles" noted by previous authors that arise when viewing and using the observed price data.

As it turns out, several of these puzzles have direct parallels in drug policy. There, hard-won experience shows that drawing conclusions from price data is not as simple as textbook economics might suggest. In Section 3 we discuss key problems observed with "naive" use of price data in drug policy and their analogies to our recent experience with price data for illicit digital goods. Of course, not every "puzzle" has a direct parallel in the drug war, but enough do to suggest experience can transfer to our research field.

In Section 4 we give a set of recommendations for future research to clarify the parallels we note here. We also debate which uses of price data may remain possible even in the case price data only conveys approximate information. Finally we discuss ways in which computer security researchers may be better positioned than drug policy researchers to obtain useful data.

In Section 5 we state conclusions. The central thesis of our paper is that observational price data is a crucial piece of security research going forward, but it must be taken with caution. Computer security is at a critical juncture with respect to public policy. Injudicious use of statistics risks giving rise to "mythical numbers" [25] which can then become enshrined in policy. Once so enshrined, these policies have far-reaching effects that are difficult to change.

2. PROMISES AND PUZZLES: PRICE DATA FOR ILLICIT DIGITAL GOODS

Promise: Better Attacker Modeling Until recently, there has been little research performed on underground markets for illicit digital goods, largely because these markets have only grown to prominence within the last decade. Before then, intrusions were motivated by a desire for fame or simple curiosity. While exchanges may have taken place of, say, passwords for compromised machines, extracting money from the activity was not the primary goal. Even today, a number of attacks take place for non-financial considerations.

Yet, in a marked change compared to the early days of computer security research, the vast majority of security attacks are financially motivated. As a result, attackers have become considerably much more predictable. Indeed, as opposed to entities motivated by ideology or other intangible goals (fame, reputation, ...), criminals driven by financial gain are going to be, by and large, economically rational. That is, they will act in their best financial interest – forgoing attacks that do not provide enough returns on the "investment" made, avoiding risky endeavors with questionable profitability, and instead, going for simpler schemes that guarantee high returns with low expected losses. As a case in point, witness the recent rise of online confidence scams [8, 21, 26].

Thus, an improved understanding of the economic environment that motivates attackers, would allow for a considerably better modeling of attacker behavior and inventives, which in turn could greatly enhance how system defenses are built. This has partially motivated some of the recent research in online crime modeling, and the associated data collection that has been undertaken.

Data Collection Moore et al. contend that online crime has become economically significant since around 2004 [21], which led to the development of markets for criminal exchange. These markets are used to exchange illicit goods such as stolen bank account credentials, botnets, credit card numbers, administrator credentials, and even full identities [13]. Participants in this economy usually negotiate sales using public forums such as websites and Internet Relay Chat (IRC), which makes it easy for researchers and law enforcement to monitor the asking prices and volume of goods [29]. Traditionally, the value of this economy has been measured in terms of total losses to consumers due to online fraud, though these estimates vary greatly. In 2008, in their annual Internet Crime Report, the Internet Crime Complaint Center (IC3) lists the total cost of online fraud at $264.6M [17]. Whereas, the U.S. Federal Trade Commission (FTC) estimates the cost of online fraud to be $1.8B in 2008 [11]. These estimates may be underestimates because they only account for fraud incidents that are reported to the respective agencies. Thus, these market valuations obviously do not include unreported incidents. But regardless of the exact value of the market, the market continues to grow because the perpetrators have a financial incentive [12].

Along with Thomas and Martin, Franklin et al. were the first to conduct a comprehensive study of the online underground economy. For eight months during 2006, Franklin et al. collected 2.4GBs of IRC logs from servers and channels that were being used to exchange illicit digital goods. They observed public offers to sell tens of thousands of credit card numbers and bank accounts totaling millions of dollars [12]. Zhuge et al. [31] describe how criminals operate in Chinese underground markets, and provide some insight regarding some of the goods exchanged in these markets by monitoring web forums where such items are advertised.

Unfortunately, this data only reflects the initial sales prices and does not include the number of completed transactions or the actual selling prices that were agreed upon after private negotiations. At the same time, naive price data promises at least two applications of great interest to security research. First, prices can in principle be used to estimate the supply of an illicit digital good, as Franklin et al. argue [12]. Second, trends in price data carry signals about the effect of an intervention in the computer security market. As we argued in the Introduction, this opens the way to a new paradigm for evaluating security research proposals.

The monitoring required for studies like these may raise legal questions. These vary from country to country and may not be clear even when a specific jurisdiction is pinned down. The cases we describe concern primarily channels and web forums that were open to the public, where the researchers were able to join the forum and then passively observe transactions. While this may be fine, research that attempts to go beyond observed prices to "testing" the market may raise serious questions. We return to this in Section 4.

Assessing the total value of online criminal markets is much more difficult than simply aggregating reported fraud losses or observed transactions. We now discuss several "puzzles" in the data that partially explain why this difficulty exists.

Puzzle 1: Steep Discounts From Face Value. The value of a financial instrument to those participating in the market is much less than it is to stakeholders outside of the market. For instance, when trafficking in online banking credentials, the primary goal of a market participant is to transfer money out of the stolen account into an account belonging to that participant. This poses significant problems in terms of performing the transaction with relative anonymity and in such a way as to not alert financial institutions or law enforcement. Thus, several intermediaries are often needed, which significantly devalues the selling price of the account credentials. As Thomas and Martin observe, a bank account with $40,000 in it may be sold for $250-500 [29]. At the same time, when the full balance is transferred out of the account, it is worth the full $40,000 to the legitimate account owner and the bank that underwrites the loss.

Furthermore, the value of the loss to the account owner may also be misleading, as the regulations require banks to absorb these costs for many types of accounts. For instance, Synovate's 2006 report on identity theft lists the mean cost of identity theft at $500 per incident, but then points out that the median cost to the account holders was $0 [28]. While this is an apples to oranges comparison, the main point is that bank business practices and U.S. law typically shield consumers from any substantial loss from an account compromise. Thus, the total value of the online criminal market changes based on the perspective from which it is being evaluated, and only until recently, the perspective has been that of U.S. consumers and financial institutions. This is important when evaluating security responses because a failure to apply the appropriate perspective will lead to skewed valuations for interventions.

Puzzle 2: Offered Prices, Not Transactions. A second puzzle is that the data collected are generally of prices offered, not of prices actually paid, which has led some researchers to question its relevance [15]. Some, but not all previous studies that have been conducted in this area [12, 29] have examined sale prices based on public offers in IRC channels. Once a buyer is interested in a particular item, she sends a private message to the seller to initiate negotiations. This is important because advertised prices are only the start of a negotiation. Thus, researchers only have data on initial asking prices, and have no idea how many transactions are actually completed or for how much buyers agree to actually pay. This is akin to reporting the average sales price of a particular vehicle by stating the sticker price, and reporting the number of vehicles sold as the total number that appear on dealers' lots!

More recent research by Kanich et al. that took over a Spam-sending botnet avoids this problem by observing the amount of spam sent and the return on investment directly [18]. While important, this research took advantage of a special opportunity to take over an active botnet. Other studies by Holz et al. and by Motoyama et al. also go beyond merely observed data. The Holz et al. study takes advantage of "dropzones," which are world-writable directories used by malware to store dumps from keylogging of compromised machines. These dropzones also contain enough information to determine the extend of compromise and amount of damage done. The Motoyama et al. study pays CAPTCHA-breaking services to break examples of CAPTCHAs from real web sites to measure the efficiency of such services. Each set of researchers lays out arguments justifying the ethical standing of their work.

In general, to measure price data from markets for illicit digital goods, researchers will need to participate in transactions. This in turn raises ethical and legal issues. We return to this in Section 4.

Puzzle 3: Lack of Standard Units or Attribution. Clayton observes that vendors in illicit markets tend to advertise "compromised bank accounts," but rarely mention specific banks [9]. Credit cards, on the other hand, may be sold based on brand or issuing country, e.g., the asking price for an American Express card is generally different from that for a Mastercard. In addition, credit cards or compromised accounts may be sold in larger or smaller lots, depending on the amount of customer service offered. The lack of a "standard unit" for trade complicates our efforts to form meaningful price series.

The lack of attribution makes it challenging to map the prices for accounts to interventions undertaken by a specific bank. In the worst case, this could lead us to mis-state the effect of an intervention. For example, if Bank of America deploys a new anti-phishing system that is perfectly effective, sellers in the market must switch to a different bank as their "supplier" of compromised accounts. If the new accounts are perfect substitutes for the Bank of America accounts, and if the cost to switch is low (e.g., rewriting a phishing e-mail and taking a screen shot of a web site), then prices may never rise. In this hypothetical scenario, the new countermeasure has worked, but the price data would erroneously tend to conclude it has not worked.

Puzzle 4: Lack of Quality Control. Another puzzle, which is relevant to all underground economies and not unique to illicit digital goods, is that there is no enforcement authority to report to when a buyer does not receive the goods as advertised. In the case of illicit digital goods, a buyer purchasing a lot of 1,000 stolen login credentials for Hotmail may find that fewer than half of the credentials are valid at the time of the purchase. Likewise, being digital goods, there is nothing to prevent a seller from selling the same set of login credentials to Alice as was sold to Bob. In these cases, when the goods are not as advertised, the buyers cannot report a seller to the Better Business Bureau or dispute the transaction with their banks. Instead, the market relies on reputations that are determined by word of mouth [12].

If an intervention is successful and a market shortage is created, a seller may opt to keep prices constant by reselling identical illicit digital goods to multiple buyers to compensate for the decreased supply. In this manner, researchers idly observing asking prices will never notice the effect of the intervention because they are unable to examine the digital goods for duplicates. Because there is rarely "honor among thieves," this scenario is likely.

3. DATA ANALOGIES TO THE DRUG WAR

We now highlight some of the data collection issues and findings from drug policy. For each issue we describe an analogy to illicit digital goods.

STRIDE Data Collection. The United States Drug Enforcement Agency publishes drug price data through a mechanism called "System to Retrieve Information from Drug Evidence" or STRIDE [30]. During investigations, undercover agents or informants will engage in actual transactions with drug dealers to obtain evidence in support of a conviction. The price paid for these drugs and the quantity received is recorded. The DEA then sends the drugs to be analyzed for purity. The resulting data has been made available to researchers since the 1980s and forms the basis for many quantitative studies of the illicit drug market.

Horowitz, however, points out several problems with using STRIDE data for policy analyses. The primary criticism is that the data are mainly gathered from buys intended to produce evidence for busts, except for a smaller program aimed solely at heroin. The price data are therefore not a uniform sample of any kind. This makes it difficult to draw statistical conclusions from the data. Furthermore, there are systematic biases in the data towards higher prices, such as the use of informants to perform buys. This causes a bias because the informants are reimbursed based on the claimed price of the drug, yielding a clear incentive to inflate the price [16].

A further issue is that the STRIDE data are not consistent with prices reported by the DC metro police. Horowitz does a analysis showing that the two agencies report a statistically significant difference in prices for heroin in the same area during the same time period. He concludes that the difference is greater than can be accounted for by normal price differences within a single city. Therefore one or both of the price data sets are not reflective of the "true" price for heroin.

Arkes et al. argue that Horowitz's claims are overstated and that the STRIDE data are still useful [2]. Specifically, they claim Horowitz improperly lumps together retail and wholesale purchases of illegal drugs. They also argue that the STRIDE data should be corrected to account for the purity of a drug bought. After adding purity to the regression and stratifying by purchase type, the supposed conflict between DEA and DC police reported prices disappears. The authors finally point out positive correlations between a decrease in STRIDE prices and decreases in emergency room visits, arguing that this is evidence of a "singal" in the STRIDE data series.

The analogy to illicit digital goods is that analyses in both settings are shaped fundamentally by the characteristics of our data gathering instrument. In the drug case, STRIDE does not attempt to be a uniform sample and its data collection was not originally designed with scientific study in mind. The controversy over STRIDE's effectiveness and the use of additional data to attempt validation of STRIDE trends highlight the need to do similar work with price data from online markets.

Experience Goods and Purity's Effect on Price Series. A key lesson learned by drug policy is that drugs are *experience goods,* meaning the quality is not known to the purchaser until after the transaction is concluded. With illicit drugs, Reuter and Caulkins report that even mid-level distributors of drugs rarely perform chemical tests of drug purity, while end users never do. As a result, distributors at each level of the drug supply chain are free to dilute the drug with additives. As a result, the product consumed by an end user is almost never 100 percent pure.

While end users cannot test the purity directly, Caulkins argues that buyers build in a dilution discount to the prices paid for illicit drugs. He proposes an*expected purity hypothesis*: the prices observed for drugs over time will depend not only on the supply, but on the buyer's expectation of the purity received. Technically, this means that regressions on price series should also regress on the purity of drugs in a given market. Caulkins et al. show that when purity is not included in a linear regression, the coefficient of price in the regression is quite low, which runs counter to what one would expect [3, 4, 7].

The analogy here with illicit digital goods is that compromised bank accounts and credit cards are also experience goods with an associated notion of "purity." For example, a buyer of bank accounts does not know before the purchase whether the account will remain open long enough to transfer money to a cashier. Credit cards have higher or lower credit limits, as well as different levels of scrutiny from issuing banks. Finally, when a buyer purchases a lot of several hundred (or thousand) credit card numbers or account credentials, it is likely that not all of these credentials will be valid; credit cards may be canceled and passwords may get changed before the transaction is completed.

While these notions of purity are not as simple as the linear scale of drug purity, anecdotal evidence strongly suggests they affect the prices buyers are willing to pay. Herley and Forencio argue that the prevalence of "rippers," participants in the market who cheat others by offering fradulent or low-quality goods, imposes a "tax" that drives down the price others are willing to pay [15].

The situation is "worse" than in the drug market, in that there is no test available for purity. For example, Herley and Florencio note that giving up the password to a compromised account to allow checking that it has the amount of money advertised also gives someone the ability to clean out the account entirely [15]. In practice, however, this may not be a material difference given the reports that drug dealers rarely test for purity even though such tests are possible.

Money Laundering. Another distinguishing factor of the illicit drug market is the use of "mules" to carry currency as payment for drug operations, instead of using traditional bank transfers. Money laundering laws in the United States and other countries require reporting transactions of certain types to law enforcement; for example in the U.S. every transaction over $10,000 must be reported. These reports allow law enforcement to trace and seize the proceeds of drug operations. As a consequence, drug organizations resort to physically smuggling currency as a method of paying suppliers and distributors [1]. Additional money laundering methods must be used to move money from drug operations to personal wealth or other businesses. Each of these methods incurs a significant cost: mules may be seized, accounts may be frozen, and moving large sums of money is difficult. As a result the amount of money that can be extracted may only be a small fraction of the money collected for illicit drugs.

The parallel here is with the use of "cashiers" in extracting liquid value from the illicit digital goods economy. Cashiers are people who transfer money from compromised online accounts to other accounts, such as those belonging to criminals. This is a high risk occupation, because the transfers create a paper trail pointing directly to the cashier. As a result, cashiers may demand a substantial percentage of the money transferred.

Price Dispersion. The price of drugs is wildly different in different cities. For example, the price of a kilogram of cocaine has consistently been $6000 more in Boston than in New York City [3]. In part, this is because Boston has less tolerance for cocaine trafficking and use than New York or other cities; this raises the cost of doing business. This phenomenon is called *price dispersion.* In a market for a legal good, we would expect these effects to disappear quickly through arbitrage. Because illicit drug markets are not allowed to advertise prices and form connections for import/export between markets in the open marketplace, however, these differences persist and can be highly lucrative for drug dealers who can establish proper connections. This is borne out by comparing price data for cocaine and heroin in different markets with price data for legal goods such as sugar and wine, then showing that the illicit drugs have higher dispersion [6].

For illicit digital goods, each IRC channel or web forum acts as a separate market. Do these markets also exhibit high price dispersion? At first we would expect this not to be the case, because after all every market is online. On

Cocaine Intervention	**Percent of Budget**	**ROI**
Domestic Enforcement	73%	$0.52
Interdiction	13%	$0.32
Source-Country Control	7%	$0.15
Treatment	7%	$7.46

Table 1: A budgetary breakdown of the four cocaine interventions that the U.S. government funded in 1992, and their respective returns on investment (ROI). Here the ROI is measured in terms of savings for every dollar spent [24].

the other hand, the markets easily discovered by the public may represent only the tip of the iceberg, while the high skill criminals trade elsewhere [15]. It is technically trivial to create markets where participation is "invite only," with invitations distributed on the basis of past successful deals.[1] We might expect the prices observed in these invite-only markets to be different, but how different, and what conclusions to draw, is not clear. For example, the number of participants in an "invite-only" market will be lower by definition than the number in an open market, leading to reduced liquidity.

Price dispersion is important because the magnitude of differences between prices in different markets determines how incomplete observation of markets affect our conclusions. If we have invite-only markets not accessible to price surveillance, yet those markets have the same prices as those which we can observe, then the existence of the invite-only markets is of less importance. On the other hand, if there is high price dispersion between different markets, and some of these markets are not accessible for price surveillance, our conclusions are weakened.

Political Solutions vs. Effective Solutions. In 1992, a total of $13B was spent in the United States on cocaine interventions. These interventions fell into four categories: domestic enforcement, interdiction, source-country control, and treatment [24]. Table 1 depicts how this budget was divided between the four interventions, and the return on investment (ROI) for each of these interventions. While domestic enforcement accounted for the majority of expenditures, the $9.5B spent on domestic enforcement mitigated societal costs of around $4.9B (i.e., every dollar spent on domestic enforcement saved fifty-two cents). Surprisingly, the least funded intervention, treatment, is over fourteen times as effective—per dollar spent—as domestic enforcement!

Given how much more cost effective drug treatment is than domestic enforcement, it is surprising that its funding pales in comparison. However, we believe that this can be explained by political considerations: in an effort to appear "tough on crime" to appease their constituents, policy makers are inclined to fund ineffective programs rather than use empirical data to optimize their decisions from an efficacy standpoint. This happens in other policy areas as well: abstinence-only sex education is repeatedly funded even though studies indicate that it does nothing to prevent teen pregnancy or to promote condom use [27].

Computer security researchers can learn from these pitfalls by evaluating interventions based on their effectiveness before policymakers make a habit of funding online computer security interventions based on political motivations. Likewise, the costs of each of these motivations needs to be evaluated holistically. While user education may yield a negative expected value for the end-user [14], banks and other stakeholders can have a positive expected value by mandating that their users undergo training [19]. From their perspective, users' time is an externality.

Now is the time to learn from these pitfalls because policymakers and law enforcement are exploring different mitigations. For example, the U.S. Federal Bureau of Investigations infiltrated a web forum used for trading illicit digital goods, used it to gather information for several years, and then arrested several of the participants. This is analogous to traditional undercover work performed for drug cases or organized crime more generally [22]. Future infiltrations may place law enforcement in a position to try market disruption tactics as suggested by Franklin et al., which undermine trust between participants in the market [12].

On the policy side, the Australian House of Representatives recently released a report suggesting that ISPs should be required to disconnect compromised machines from the network [23]. Scott Charney, corporate VP of Trustworthy Computing at Microsoft, suggested a similar approach in his keynote at the RSA Conference 2010 [20]. We do not see a clear analogy here with interventions in the drug war, but the debate surrounding this intervention needs accurate data on how effective such a measure would be at stopping computer crime.

4. HOW CAN SECURITY RESEARCH LEARN LESSONS FROM DRUGS?

We have highlighted four specific places where experience from drug policy is directly analogous to open issues in using price data for illicit digital goods. What do we do with these analogies? Where should we go next? We now make several suggestions for future research.

First, we have just scratched the surface of parallels between drug markets and illicit digital goods. While we have argued primarily the similarities here, the markets have clear and important differences. Most obvious is the simple fact that drugs are physical goods, which have non zero marginal cost to produce and must be smuggled into a country by land, air, or sea. We suggest a more in depth study of these similarities and differences to uncover future opportunities for knowledge transfer between the two communities.

Second, we suggest investigation into how the security community can best craft a data gathering instrument to further research. This investigation will need to include legal and ethical issues as well as traditional computer science or statistical issues. Law enforcement may already engage in transactions in markets for illicit digital goods, in which case a system analogous to STRIDE for recording the prices paid may be a reasonable first step. Such a system, of course, would have many of the same difficulties as STRIDE with respect to non uniform sampling.

The computer security community is in some respects better positioned to create data gathering instruments than the drug policy community.[2] Gathering price data rarely incurs personal risk of violence or injury, unlike with drug price data. Transactions that take place online can be captured

[1] The authors recall bulletin board systems in the late 80s which had such invite-only policies for particularly juicy warez collections.

[2] We thank Jason Franklin for pointing this out.

in their entirety for other researchers to re-visit with near zero cost.

On the other hand, research into price data raises thorny legal issues, especially if we want to go beyond advertised prices. Is it legal for security researchers to actually carry out transactions with criminals to learn the true price of a good or service? What if a researcher advertises an illicit good solely for the purpose of soliciting bids, with no intention of completing the transaction? Do the answers change if the researcher is working with a bank to focus specifically on accounts belonging to that bank? Do the answers change if the researcher works with law enforcement? What if the criminals under study are in a different country than the researchers? The answers likely depend on the jurisdiction and the situation, but we need to develop clarity here to prevent unfortunate incidents where researchers end up in legal trouble while investigating markets for illicit digital goods.

Third, we suggest reaching out to economists and statisticians find ways of tolerating noisy or incomplete data in conclusions. While in principle all data is useful, in practice drawing incorrect conclusions due to data problems or unfounded application of statistical techniques is costly. Put another way, bad data is sometimes *worse than no data at all*, because it leads us to commit resources best spent elsewhere or make irrevocable decisions about the design of our systems. How can we design mechanisms that limit this downside risk from bad data?

Finally, we suggest finding ways to directly integrate knowledge obtained from price surveillance into interventions with a tight "feedback loop." The goal here is to quickly detect spurious inferences that may arise from imperfections in our data gathering and analysis. For example, consider a CAPTCHA that is used to protect signups for a service. Many CAPTCHAs can be tuned to make it more difficult or less difficult for adversaries to defeat them, at the cost of making it more difficult for legitimate users to pass. We could imagine a mechanism that picks the value of the tradeoff based on the best price data observed for accounts of that service. If the price drops, this may indicate that the adversary has adapted to the current difficulty of the CAPTCHA and can break it more easily than before. The defender can then react by increasing the difficulty of the CAPTCHA or instituting a new CAPTCHA altogether. Because the price data in this case can be cross-checked against other measures of fradulent activity on the service, an incorrect decision based on the price data to make the CAPTCHA more difficult may conceivably be detected and corrected.

5. CONCLUSION

Let us be clear: we want observational data in security research. We look forward to the day when squabbling about the arcane details of econometric technique is the largest of our worries. Looking at the growing work on observational economics of online criminal markets, we can see many parallels with the drug trade. For instance, both have producers who do not interact directly with their consumers, both deal in economies of scale, both see products go through many hands to complete a transaction, etc. We have argued that research on data collection and use in the illicit drug market may yield valuable insight into how to intervene in the online criminal market, or at least in how to avoid pitfalls in evaluating interventions.

As a parting motivation, we note interventions in the drug market are not purely rational because politics may trump practical considerations. For example, a U.S. policymaker may be more likely to support stricter prison sentences than treatment programs, because that makes the policymaker appear to be "tough on crime" to constituents, despite the fact that treatment programs are cheaper and more effective. Interventions in the online criminal market are newer and may suffer from less received wisdom. The challenge for security research going forward is to learn its lessons from the "war on drugs" and implement these lessons before it is too late.

6. REFERENCES

[1] U. D. E. Agency. Inside the DEA, DEA programs, money laundering, 2010. `http://www.justice.gov/dea/programs/money.htm`.

[2] J. Arkes, R. L. Pacula, S. M. Paddock, J. P. Caulkins, and P. Reuter. Why the dea stride data are still useful for understanding drug markets. NBER Working Papers 14224, National Bureau of Economic Research, Inc, Aug. 2008.

[3] J. P. Caulkins. Price and purity analysis for illicit drug: Data and conceptual issues. *Drug and Alcohol Dependence*, 90(Supplement 1):S61 – S68, 2007. Behavioral and Economic Perspectives in Drug Abuse Research.

[4] J. P. Caulkins and R. Padman. Quantity discounts and quality premia for illicit drugs. *Journal of the American Statistical Association*, 88(423):748–757, 1993.

[5] J. P. Caulkins and P. Reuter. What price data tell us about drug markets. *Journal of Drug Issues*, 28:593–612, 1998.

[6] J. P. Caulkins and P. Reuter. Ilegal "lemons": price dispersion in cocaine and heroin markets. *Bulletin on Narcotics*, LVI(1-2):141–165, 2006.

[7] J. P. Caulkins and P. Reuter. Illicit drug markets and economic irregularities. *Socio-Economic Planning Sciences*, 40(1):1–14, March 2006.

[8] N. Christin, S. Yanagihara, and K. Kamataki. Dissecting one click frauds. Technical Report CMU-CyLab-10-011, Carnegie Mellon University, Apr. 2010.

[9] R. Clayton. Re: street prices for digital goods?, November 2008. http://www.mail-archive.com/cryptography@metzdowd.com/msg09835.html.

[10] DEA Public Affairs. New data show significant disruptions in U.S. methamphetamine, cocaine markets price of meth soars 73 percent; purity down by nearly a third, November 2007.

[11] Federal Trade Commission. Consumer sentinel network data book for january-december 2008. Technical report, Federal Trade Commission, 2009.

[12] J. Franklin, V. Paxson, A. Perrig, and S. Savage. An inquiry into the nature and causes of the wealth of internet miscreants. In *CCS '07: Proceedings of the 14th ACM conference on Computer and communications security*, pages 375–388, New York, NY, USA, 2007. ACM.

[13] D. Geer and D. Conway. What we got for christmas. *IEEE Security & Privacy*, page 88, January/February 2008.

[14] C. Herley. So long, and no thanks for the externalities: the rational rejection of security advice by users. In *NSPW '09: Proceedings of The 2009 New Security Paradigms Workshop*, pages 133–144, New York, NY, USA, 2009. ACM.

[15] C. Herley and D. Florêncio. Nobody sells gold for the price of silver: Dishonesty, uncertainty and the underground economy. In *WEIS '09: Proceedings of the 2009 workshop on Economics and Information Security*. ACM, 2009.

[16] J. L. Horowitz. Should the DEA's STRIDE Data Be Used for Economic Analyses of Markets for Illegal Drugs? *Journal of the American Statistical Association*, 96(456):1254–1262, December 2001.

[17] Internet Crime Complaint Center. 2008 internet crime report. Technical report, Bureau of Justice Assistance, 2008.

[18] C. Kanich, C. Kreibich, K. Levchenko, B. Enright, G. Voelker, V. Paxson, and S. Savage. Spamalytics: An empirical analysis of spam marketing conversion. In *Proceedings of the Conference on Computer and Communications Security (CCS)*, Alexandria, VA, Oct. 2008.

[19] P. Kumaraguru, J. Cranshaw, A. Acquisti, L. Cranor, J. Hong, M. A. Blair, and T. Pham. School of phish: a real-world evaluation of anti-phishing training. In *SOUPS '09: Proceedings of the 5th Symposium on Usable Privacy and Security*, pages 1–12, New York, NY, USA, 2009. ACM.

[20] E. Mills. Microsoft exec: Infected PCs should be quarantined (Q A). `http://news.cnet.com/8301-27080_3-10462649-245.html`.

[21] T. Moore, R. Clayton, and R. Anderson. The economics of online crime. *Journal of Economic Perspectives*, 23(3):3–20, Summer 2009.

[22] U. F. B. of Investigation. Dark market takedown: Exclusive cyber club for crooks exposed, 2008. `http://www.fbi.gov/page2/oct08/darkmarket_102008.html`.

[23] A. H. of Representatives Standing Committee on Communication. Hackers, fraudsters and botnets: Tackling the problem of cyber crime, 2010. `http://www.aph.gov.au/house/committee/coms/cybercrime/report.htm`.

[24] C. P. Rydell and S. S. Everingham. *Controlling cocaine: supply versus demand programs*. Rand, Santa Monica, CA, 1994.

[25] M. Singer. The vitality of mythical numbers. *Public Interest*, 23(1):3–9, 1971.

[26] F. Stajano and P. Wilson. Understanding scam victims: Seven principles for systems security. Technical Report UCAM-CL-TR-754, Cambridge University, Aug. 2009. To appear in Communications of the ACM.

[27] L. S. Stepp. Study Casts Doubt on Abstinence-Only Programs. *The Washington Post*, April 14 2007. http://www.washingtonpost.com/wp-dyn/content/article/2007/04/13/AR2007041301003.html.

[28] Synovate. Federal Trade Commission - 2006 Identity Theft Survey Report. Technical report, Federal Trade Commission, 2007.

[29] R. Thomas and J. Martin. The underground economy: Priceless. *USENIX ;login:*, December 2006.

[30] U.S. Department of Justice. DEA Major Information Systems, December 1998. http://www.justice.gov/dea/foia/stride.html.

[31] J. Zhuge, T. Holz, C. Song, J. Guo, X. Han, and W. Zou. Studying malicious websites and the underground economy on the Chinese web. In *Proceedings (online) of the Seventh Workshop on the Economics of Information Security (WEIS)*, Hanover, NH, June 2008.