Beneath the radar: Exploring the economics of business fraud via underground markets

Carlos H. Gañán
Delft University of Technology
c.hernandezganan@tudelft.nl

Ugur Akyazi
Delft University of Technology
u.akyazi@tudelft.nl

Elena Tsvetkova
Delft University of Technology
e.o.tsvetkova@student.tudelft.nl

***Abstract*—Underground marketplaces have emerged as a common channel for criminals to offer their products and services. A portion of these product comprises the illegal trading of consumer products such as vouchers, coupons, and loyalty program accounts that are later used to commit business fraud. Despite its well-known existence, the impact of this type of business fraud has not been analyzed in depth before. By leveraging longitudinal data from 8 major underground markets from 2011-2017, we identify, classify and quantify different types of business fraud to then analyze the characteristics of the companies who suffered from them. Moreover, we investigate factors that influence the impact of business fraud on these companies. Our models show that cybercriminals prefer selling products of well-established companies, while smaller companies appear to suffer higher revenue losses. Stolen accounts are the most transacted items, while pirated software together with loyalty programs create the heaviest revenue losses. The estimated criminal revenues are relatively low, at under $600 000 in total for the whole period; but the total estimated revenue losses are up to $7.5 millions.**

I. Introduction

Over the past decades, industry reports revealing the costs of cybercrime are becoming more and more common, with values ranging from a few millions to trillions of dollars [1]–[3]. While the accuracy and validity of these reports have been widely questioned in the academic literature [4], [5], they all have something in common: cybercrime has a significant economic impact on society. Nevertheless, the lack of empirical evidence for certain types of cybercrime has swayed these reports over specific types of illegal activities that are more visible and covered by the media (e.g., ransomware attacks [6]–[8]).

Trendy and novel attacks keep catching the attention of journalists and researchers, while the impact of other types of fraud gets belittled. In this paper, we focus on business fraud, that is fraud committed by consumers after the acquisition of illegally obtained items, such as vouchers, gift cards, coupons, pirated software, stolen accounts and loyalty rewards. The sparsity and underreporting of this type of fraud makes it especially hard to quantify *ex post*. Thus we investigate different types of business fraud before they are materialized by analyzing how criminals trade illegal consumer products in underground markets.

Underground markets constitute a commercial channel for miscreants to sell illegal products and services, but at the same time, they also present unique opportunities to quantify illicit activities that might be underrepresented in police reports. These markets have been steadily evolving over the last few years [9]. They are no longer the chaotic and poorly regulated chat forums of the past, but rather sophisticated platforms turning criminal activity into a commodity, easily accessible by consumers. While the largest amount of goods traded on these markets are drugs, the exchange of stolen financial information, identity theft, fake accounts, as well as various types of crimeware also have a stable share [10]. A smaller portion of these goods can be assigned to a category comprising miscellaneous fraudulent items: stolen accounts, vouchers, coupons, gift cards, loyalty program rewards.

Previous studies have focused on the profits made by the criminals facilitating digital products from business-to-business [10], [11] but neglected the trades in these markets targeting directly consumers. On the other hand, a handful number of qualitative studies focused on understanding criminals and victims of particular types of fraud without quantifying its financial impact (e.g., airline ticketing fraud [12]). In this paper, we analyze the effect on the companies suffering from this illegal sell of consumer products in underground markets. The goal of this study is to understand the impact of business fraud facilitated via anonymous marketplaces, which has remained unexplored, and draw conclusions about the factors and characteristics which make a company an attractive target for criminals. We adopt the FBI's definition of business fraud, i.e., "dishonest and illegal activities perpetrated by individuals or companies in order to provide an advantageous financial outcome to those persons or establishments".

To reach this goal, we analyze scraped data on the business-to-consumer products of 8 anonymous marketplaces consisting of 29 936 product listings and over 300 000 transactions. Only 2 318 listings corresponded to business fraud while the rest were related to other cybercriminal services. We estimate the official pricing of the products being sold in the anonymous markets by manually collecting publicly available data on these products. Our models show that pirated software together with loyalty program fraud cause the largest revenue losses, while account fraud makes up the greatest part of all transactions which have taken place. In general, most products retail for relatively low prices compared to their official price. In turn, account fraud caused the highest revenue losses, due to the large number of points stored in the stolen accounts.

To further understand why these businesses suffered from business fraud, we gather additional data related to the characteristics of the affected companies. Based on Routine Activity Theory (RAT) [13], we identify characteristics that help to

understand the criminals' target selection process (i.e., size, popularity, reputation and area of service). We use these characteristics to explain the differences on revenue losses across the affected companies. The results show that popularity significantly influences the number of fraud products being sold; while company size together with a wider area of service impact a company's revenue losses.

In detail, our contributions are the following:

1) We present the first quantitative analysis of consumer-driven business fraud facilitated via online underground markets, covering 8 markets over 7 years;
2) We classify the different types of business fraud depending on the illegal products being sold in the underground markets, i.e., stolen accounts, vouchers, pirated software and loyalty programs rewards;
3) We identify 47 companies whose products were illegally being sold and estimated their revenue losses;
4) We identify and measure several factors related to the visibility and value of a company to understand why some companies suffer more revenue losses than others.

II. Related work

A. Cost of cybercrime

Estimating the cost of cybercrime has been a recurrent topic of discussion among scholars. Numerous reports by various parties, from governmental organizations to security vendors, have aimed at estimating the costs of cybercrime. In the late 90s, the FBI launched the Computer Crime and Security Survey which estimated the cost of computer crime in several billion dollars [14]. Since then, a myriad of reports have appeared; some just looking at specific types of cybercrime (e.g., cost of malware [15]), others analyzing costs at the country level (e.g., US [16], Belgium [17]) and another set looking at the societal costs (e.g., long-term costs [4]).

However, oftentimes the presented analyses are based on data that is either over- or under-estimated [18], either intentionally or due to the measurement techniques employed, as explored by Anderson *et al.* [19]. This presents a framework that identifies direct, indirect and defense costs. Rather than adding up the partial estimates, they argue it is more informative to present these impacts separately.

Only a handful studies provide parts of a comprehensive assessment on the cost of cybercrime. Some of them focus on articulating a model to enumerate the different impacts (e.g., [20], [21]), while others limit their scope to specific data set, such as data breaches (e.g., [22], [23]) or losses related to credit card fraud [24]. Even smaller is the number of studies that try to empirically estimate the cost of cybercrime by: (i) categorizing different cost types; and (ii) using data from diverse sources to estimate costs, either at the national or at the global level. For instance, McAfee reports [25] and Detica [26] are an example of these.

B. Underground Digital Trade

The digital underground economy is constantly evolving and changing, thus making it difficult to get a comprehensive view of the way it is organized. What started as forums used primarily for the purpose of sharing experience and techniques by hackers and individuals interested in honing their skills, has become a complex network mostly driven by profit [27].

Online illegal trade has existed in one form or another for several decades. Hackers and other interested parties are reported to have been communicating and sharing files through message and bulletin boards as early as the 1980s [28]. Such endeavors were made easier by the emergence and later the global spread of the Internet, which in turn led to the creation of better organized networks and criminal groups. Thomas et al. [29] describe the initial actual digital underground markets as platforms utilizing IRC: a standard protocol for real-time text messaging over the Internet. Users on the illicit markets would utilize IRC to share availability and pricing information about various products and services, such as credit card information, compromised accounts, botnets, malware.

Gradually, this type of structure was replaced by web forums, where more varied information was shared under unique threads, and access was often more restricted than in the open IRC chats [30]. The existence of such forums in various geographical regions has been explored previously: Holt *et al.* [31] present an analysis of markets operating in Russian and English, Broseus *et al.* [32] explore markets in Canada, while Yip *et al.* [33] focus on forums in China.

Eventually, webforums evolved into a market structure, similar to e-commerce platforms popular on the open web. The predecessor of the numerous underground markets which have existed in the last decade and the first to boast an organization of this kind was the Silk Road market, launched in 2011 [34]. It is reported that by the time it was shut down in law enforcement operations in the US and Australia in 2013, a revenue of around $1.2 billion had been made through the site [35]. Its closure, however, did not lead to a decline in the underground trade online, as several other markets of a similar structure emerged, drawing in the users of the Silk Road [28].

Despite the relatively recent creation and following closure of the Silk Road, the underground economy has been flourishing. Buxton and Bingham [36], Soska and Christin [9], Broseus *et al.* [32] argue that there has been an increase in the volumes of goods and services flowing through the crypto markets as well as in the range of products offered. This could be credited to the developments in technology, making participation in the illegal trade more secure.

Soska and Christin [9] showed that since the existence and disappearance of the first online anonymous market, the number of sellers has significantly increased along with the high competitiveness among suppliers. However, most of the examined markets seem to reach vendor saturation, or never expand sufficiently, due to law enforcement operations shutting them down; self-destructing mechanisms such as exit scams performed by the markets' owners; or voluntary closures [37].

Despite the limited growth of the markets and their short life span, the activity on them remains high as shown by Broseus *et al.* [32]. Their findings are consistent with the previously reported [9] diversification and replication of vendors on dif-

ferent marketplaces, aiming to increase profits and reputation or mitigate risks of potential shut-downs. The study serves to confirm what other authors have observed, despite its regional character owing to the fact that the analysis is based only on data pertaining to markets operating from Canada.

Although law enforcement operations are seen as successful in closing down certain markets, they have actually led to the emergence of even more dark net markets, utilizing more sophisticated technologies and spurring innovations in the process. As witnessed in 2017, when three of the most significant global dark web markets: Alphabay, Hansa, and RAMP were shut down in law enforcement operations, these actions only led to the trade being shifted to other existing markets or to newly-found smaller, privately run vendor shops, along with more regional secondary markets operating in particular countries or languages [37]. It is expected that the smaller scale and more targeted approach of these marketplaces is possibly not going to attract the same level of attention from the authorities or the media, as the larger platforms.

III. Data collection and method

To estimate the impact of business fraud on the affected companies, we looked into 3 different components: (i) number of transactions that were carried out in the underground markets; (ii) retails prices of products involved in business fraud; and (iii) characteristics of the affected companies.

A. Business fraud via underground markets

We use secondary data sources to explore that extent to which underground markets facilitate business fraud. In particular, we leverage scraped data from online anonymous markets made available via the IMPACT project [38]. This dataset includes records collected from eight prominent underground anonymous marketplaces (Agora, Alphabay, Black Market Reloaded, Evolution, Hydra, Pandora, Silk Road 1 and Silk Road 2) from 2011 to May 2017, and consists of 44 671 listings and 564 204 transactions made on digital goods, grouped in 17 categories. Unfortunately, these categories did not distinguish products related to business fraud which are dispersed across multiple categories.

1) Extracting business-fraud products: All eight examined markets operate mostly in the English language, and were active during various time periods. The dataset contains transaction information starting from the original Silk Road 1, and follows its successors, including one of the largest existing single marketplaces to date: Alphabay. While the trade on the markets revolves predominantly around drugs and similar substances, a portion constitutes digital goods, such as malware, bots, fake and pirated goods, as well as various types of information. Cybercriminal listings in the dataset have previously been classified [10] into several categories, differentiated by whether they are aimed at other criminals, as in business-to-business, or to be used personally. On the contrary, we focus on the retail side of the trade.

While some of the categories in the dataset already could be discarded as unrelated to business fraud, we had to conduct some additional filter over some generic categories which contained a mixed variety of products. The original dataset is grouped in the following categories: `Accounts`, `Fake`, `Guide`, `Pirated`, and `Voucher`, along with `Custom` and `Others`. An overview of the categories is provided in Table I, along with the original number of listings in each category and the remaining number after a filtering process.

TABLE I
BUSINESS FRAUD-RELATED CATEGORIES

Category	Description	Total #Listings	Business-fraud #Listings
Accounts	Accounts for media streaming services; pornography websites	3 805	1 071
Fake	Fake items: ID cards, passports, money	3 429	23
Guide	Guides for various illegal endeavors; hacking; personal development	5 097	63
Pirated	Pirated software and other digital content	1 431	148
Voucher	Vouchers and gift cards for various stores and restaurants; lottery tickets	1 305	762
Custom	One-time buyer- specific products or services	6 378	28
Other	Miscellaneous items: clothing and accessories replicas; drugs; documents; cash	8 491	223

Our filtering process aimed at distinguishing business fraud related items. We inspected manually the data from the categories outlined in Table I as the nature of the content of the listings makes it unfeasible to automate the task. The text in the listings included various jargon terms, abbreviations, unrelated keywords, and other incomprehensible phrasing that would reduce the accuracy of any automated classifier. Out of the original 29 918 records, 2 318 contained business-related fraud listing. The other listings were excluded as irrelevant to the research topic for one of the following reasons:

- listings concerning drug offerings which had slipped through in the digital goods categories during the initial categorization of the dataset;
- incomplete listings, where it is unclear what is the offered product, or which is the concerned company;
- pornography accounts, as they were considered out of the scope of this research for ethical concerns;
- unrelated listings: invitations for underground markets, lottery tickets.

As we are interested in business fraud, we expected the categories of `Accounts` and `Voucher` to be central to the subject, yielding 1 883 records. However, we also inspected the remaining categories for entries which could be relevant, which resulted in the addition of 485 more records.

2) Classifying business-fraud products: The next step of the data preparation process entailed categorizing the selected listings depending on the type of the related product or service. This resulted in four categories (see Table II): `Accounts`, `Loyalty programs`, `Vouchers`, and `Pirated software`.

TABLE II
NUMBER OF LISTING, DEFRAUDED COMPANIES AND TRANSACTIONS PER FRAUD CATEGORY

Category	Total # listings	Business fraud # listings	Total # companies	Consumer Defrauded # companies	# Transactions
Accounts	1 067	711 (66.6%)	121	22	30 038
Vouchers	887	524 (59.1%)	196	16	8 816
Pirated Software	317	288 (90.9%)	23	7	7 121
Loyalty programs	154	94 (61%)	29	7	800

Then, we extracted the company or business which was affected for each listing. Based on this classification, we estimated the amount of sales per listing per company by looking at the feedbacks left by the buyers; hence providing data for the number of transactions for each listing, and the exchanged amount. As there are no official records or statistics about the volume of sales transpiring on these marketplaces, we used the number of feedbacks as a relatively reliable measure for sales [9], [10]. Since markets require customers to leave a feedback score after making a purchase, these feedbacks correspond to the lowest number of transactions which have taken place.

Finally, we anonymized the name of the companies whose products were being illegally sold in the anonymous markets. As some of the information concerning the affected companies can be considered sensitive, we replaced the name of each company with an alphanumerical code, based on the category the firm was placed in.

B. Product's retail prices

To examine the impact of business fraud facilitated via anonymous markets, we gather additional information related to the retail prices at which the products could be bought legally. To the best of our knowledge, it does not exist a database of prices per product, so we gathered manually this data by exploring the websites of the various selected companies and extracting the product retail price.

We followed a systematic process for determining the retail selling prices of the various products:

- For service accounts and subscriptions as well as software subscriptions, the official website usually supplied several options ranging from subscribing for various time periods (monthly, quarterly-, semi- or annually), to signing up for basic, or premium services. In order to avoid over-estimations, we selected the price for the most basic service offered, at the most favorable terms, which is usually the annual subscription. Considering most items offered on the underground markets are listed as lifetime accounts, we annualized the retail prices for the service, not accounting for their value in perpetuity.
- In the case of loyalty programs, the offers on the dark web listed the amount of points accumulated in the accounts. Thereby, we examined the website of the respective company so as to determine the monetary equivalent of a point, and subsequently calculate the actual value of the account on offer. When this information was not made officially available, or the access to it was restricted to only members of the service, secondary sources were consulted, such as other websites providing relevant information; and FAQ (Frequently asked questions) sections.
- In the case of vouchers and coupons, the description of the listings state a value for the discount they offer stated in the description, which was the one used for the analysis. In the case the discount or the gift card was in the form of a free item, we consulted the respective website of the company produced the item so as to estimate the worth of the offered item, again selecting the most affordable option.

C. Company characteristics

We gathered additional information to characterize the properties of the affected companies. Based on the Routine Activity theory, this data aimed at characterizing the value, visibility and accessibility of a company. 'Thus, we estimated the following factors:

- *Company size*: as there is no open-source dataset providing financial information for all companies and all indicators, we gathered this information from officially published annual reports and financial statements for the examined period from the companies' corporate websites or government institutions. In the case such were not public, we consulted third party sources such as marketing databases. In particular we collected revenues, net income, total assets, total equity, and number of employees for each company.
- *Domain popularity ranking*: this information serves to define the Visibility aspect, as it reflects the online presence of a company. We used the Cisco Umbrella Popularity List[1], which consists of the top 1 million most popular internet domains, measured by the number of unique client IPs visiting a domain relative to the sum of all domain requests. While the list is updated regularly, it did not exist at the moment some listing were posted; for those companies we used the first Umbrella Popularity List from 2016. For the rest of the companies for which products were sold in 2016 and 2017, we computed the average ranking from the list collected from 2016 to 2017.
- *Reputation*: we estimated the reputation of a company based on the Global Top 500 Companies list [2], which is a popular measure to evaluate companies' corporate reputation. Again, we computed the average ranking based on

[1]https://umbrella.cisco.com/blog/cisco-umbrella-1-million
[2]https://brandirectory.com/rankings/global/2017/

the different lists and the years when the products were sold in the anonymous marketplaces.

- *Area of Service*: this information specifies whether the examined companies provide their services on a global or a local level: for instance, only in a specific country or a geographic area. We collected this information from different data sources: preferably the company's corporate website, or when unavailable: a third source such as marketing databases, and is traced back to the year 2017.

IV. Quantifying Consumer-Based Business Fraud

Business fraud consists of illegal activities perpetrated by individuals or companies in order to obtain an advantageous financial outcome [39]. For the purpose of this paper, we focus on business fraud that is facilitated by consumer products being sold in underground markets. To quantify its impact, we analyze 3 different dimensions: (i) the revenue made by criminals selling these products; (ii) the number of sales and vendors offering these products; and (iii) the potential revenue losses incurred by the companies whose products are being illegally sold.

A. Criminals' Revenue

To quantify how much money criminals made by facilitating business fraud, we analyzed the scraped underground market data and estimated the criminal's revenue for each of the identified companies. We break down this analysis per product category.

1) Stolen Accounts: A total of 22 companies had accounts being sold in the underground markets. The majority of the targeted companies (15 companies) are media services providers, such as video and music streaming, telecommunications and sports. Among the affected companies there is also a transportation company, an educational service, a video game, and 4 e-commerce websites.

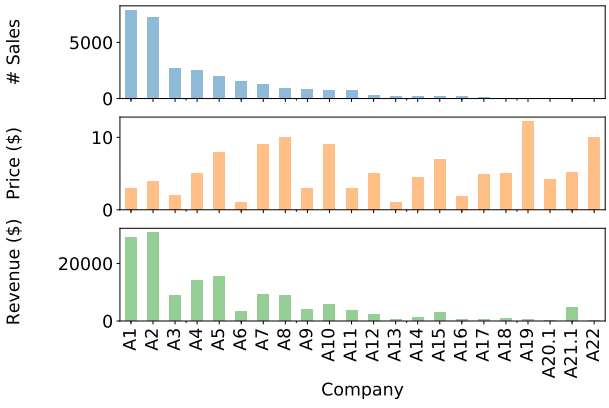


Fig. 1. Revenue, Sales, and Price per Company in "Accounts" category

There are two companies which created 43% of the total revenue of the criminals transacting account information. These companies are popular music and video streaming providers, from which the criminals made $30 674 and $29 141 (Figure 1). This is nearly twice as much as the next two entries in the list: the telecommunications companies A5 and A4, amassing $15 325 and $14 139, in turn. The rest of the companies on the list have led to less than ten thousand dollars of revenue for their sellers.

Furthermore, Figure 1 shows that there is no apparent relation between the median price per transaction and the total amount of money collected from sales for each company affected. The high values for A2 and A1 are due to the large amount of sales made on the accounts for both, as the price they were sold for is on average less than most others in the category: $3.94 for A2, and $2.99 for A1. However, the sales of such accounts are considerably more than for any other service across all categories, with more than 7 000 transactions made for each. This could be explained both with the ubiquitous use of the two services, leading to their high demand in the retail underground trade, as well as the wide availability of potential accounts for hijacking.

Accounts for media services providers and video streaming (A1, A2, A3, A4, A5, A7), especially for sports services (A8, A9, A10, and A12), are among the most widespread items for sale, while accounts for e-commerce websites and online retailers are the least offered (A19, A20.1, A22). Retail accounts usually include previously accumulated amounts to be spend or other financial information, such as credit card details, which could explain the prices they fetch on the dark web markets of around $10, considering these services are officially offered for free. This is especially the case with A21.1, which has made an estimated $4 705 from merely 33 sales.

We next examine the evolution of criminals' revenue over time. Figure 2 shows that items from the accounts category were generally offered throughout the larger part of the examined period. Clothing retail stores accounts are an exception to this, such as A22, and A19, which were available for a short interval of a month or two. Similarly, the accounts for A21 were sold for a limited time, though the amounts they made for a single listing in 2014 were substantial.

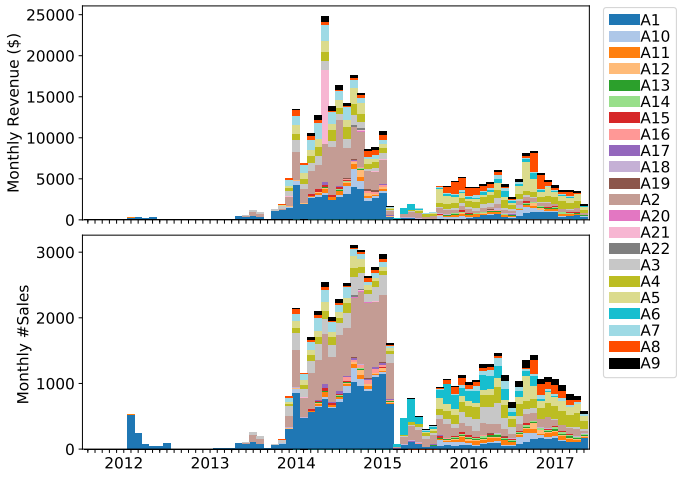


Fig. 2. Number of sales and criminal revenue due to accounts selling

As can be expected, the revenue is strongly dependent on the functioning of the underground markets themselves. The

lower values between 2011 and 2013 represent entirely the offering on the first anonymous market Silk Road 1, in the initial stages of the underground trade, and the slight decrease when it was shut down. The following steep rise observed from 2014 to 2015, marks the highest point in sales over the whole period for all the categories, except loyalty programs, which have their peak in the summer of 2016. This increase coincides with the appearance of several new markets: Evolution, Silk Road 2, Hydra, and Agora. Furthermore, the following plunge in the beginning of 2015 is closely related to the disappearance of these markets for various reasons from police take- downs to exit scams. The subsequent gradual rise in generated revenues is due to the emergence of the Alphabay market, which during its existence until mid- 2017, drew in considerable traffic.

2) Loyalty Programs Accounts: Compared to the accounts category, there are less affected companies in the loyalty programs category, i.e., only 7 companies. Four of these are commercial airlines (L1, L4, L5 and L7), whereas the remaining three are hospitality companies (L2, L3, L6). The number of loyalty program accounts is relatively low compared to the other categories, and so are the number of listings and revenue made from them.

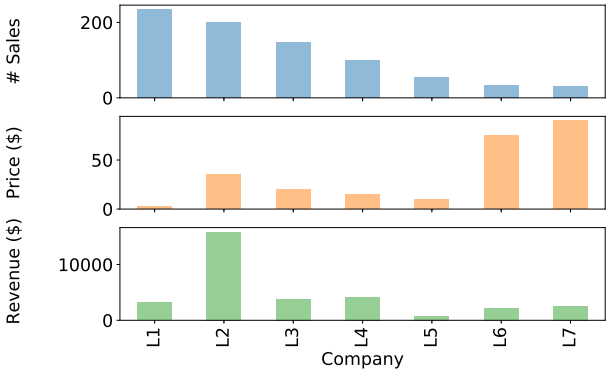


Fig. 3. Revenue, sales, and price per Company in "Loyalty Programs" category

The most targeted company in the category is the airline L1, and its frequent-flyer program, which has registered 234 sales, made on 18 listings. However, airline fraud is far from being the most lucrative on the underground markets, having brought only $3 220, which is considerably less than the highest criminal revenue, made from exploiting the hotel chain L2: $15 818. This is due to the low value of L1 accounts: their median value is $2.5, the lowest in the category. On the other hand, L2 accounts have a median value of $35, and one of the highest averages of nearly $80. The only low-cost airline featured in the category: L5, has a few accounts, which are being offered for prices lower than the rest, and consequently has accumulated the smallest amount of money: an estimated $787.

Figure 3 shows the total revenue made on the examined underground markets, and the median prices for each of the selected companies. The amount of money made is primarily related to the number of transactions, with the top four companies which have the most listings being at the upper half of the list. The two least common accounts: those of the airline L7, and the hotel chain L6, have achieved similar revenues to the top earners despite their low levels of sales, owing to their very high price per product. There have been around 30 transactions made for each, though at median values as high as $90, and $74.91, respectively.

Furthermore, unlike other products traded in the other categories, none of the loyalty programs accounts have been sold for free, all having positive minimal values. They have also consistently achieved high maximum amounts, of a few hundred dollars each.

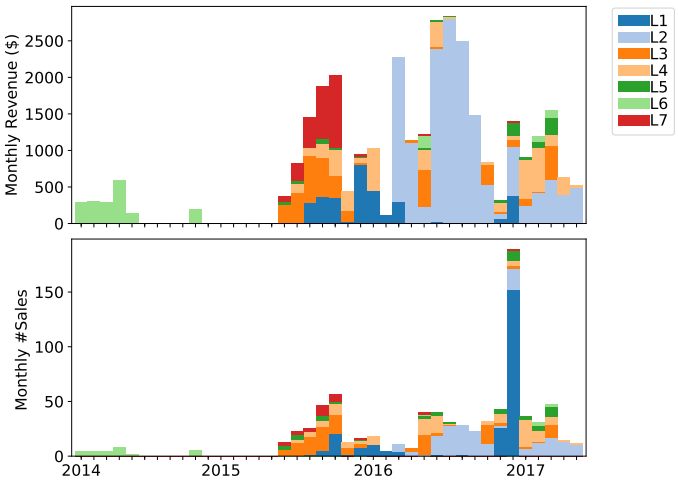


Fig. 4. Number of sales and criminal revenue due to loyalty programs

When examining the evolution of loyalty programs fraud over time (see Figure 4), it is worth noting that loyalty program accounts were offered more sporadically than other products, for a few months at a time, likely after data leaks have occurred. The hotelier's accounts were also the most consistently sold on the markets: from 2016 until the end of the examined period.

Interestingly, there are no records for such fraud before 2014, when L6 accounts were first offered for sale. The peak in the revenues made from loyalty accounts is in 2016, after a bulk of high-value L2 accounts were traded on Alphabay. This is most probably related to a large data breach reported by the same company in 2015, when more than half of the company's locations were impacted, leading to the exposure of their customers' personal information and payment card data. Moreover, the highest point in sales of loyalty accounts is also in 2016, although it is linked to the increased offering of L1 and L4 accounts, which sold for less than those of L2.

3) Pirated Software Fraud: Similar to loyalty programs, there was a modest revenue in the pirated software category, despite that the number of offerings was higher. A wide variety of products were sold on the underground markets, though the greater part were issued by a few companies, for instance, the listings for S2 include offers for their several products. For this reason, the offers for the various products are grouped under their corresponding brand. The listings for S1.1 and S1.2 are the only ones listed as separate entities, due to their different

nature. The two products are also the most widely traded on the markets, and along with the S2 software significantly outperform the rest in the category by sales and accumulated revenue.

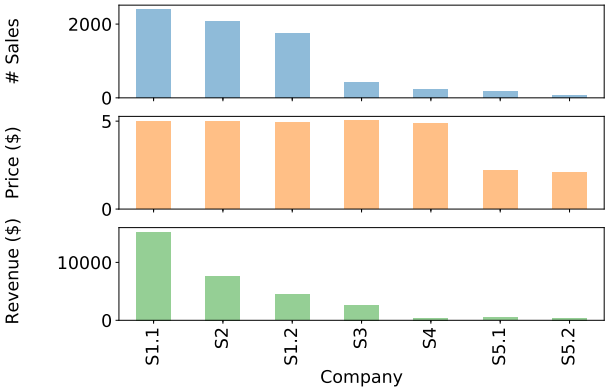


Fig. 5. Revenue, Sales, and Price per Company in "Pirated Software" category

There are very slight differences between the prices of these products: most of the products were sold for a price of around $5. The exceptions are the listings for the antivirus software S5.1, which sold for $2.2, and the computer tool S5.2, owned by the same corporation, retailing for $2.06. Because of this uniformity, the differences in the money made on account of each company are strongly dependent on the number of sales made for each. Thereby, S1.1 takes the lead with 2 394 successful transactions, contributing to $15 242 of revenue for its vendors. The second most targeted company, S2, made approximately half that revenue with $7 692 from 1 403 transactions, while the third product, S1.2, accumulated a little under $4 500 from 784 sales. The fourth most targeted item in the list: the educational software S3, was sold nearly 500 times, but due to its relatively high price made $2 695.

Pirated software products (see Figure 6) were traded continuously over the span of the whole period. The exception to this are items of the educational software S3, which appeared solely during the peak of the trade in 2014-2015. Additionally, S1 and S2 products seem to hold a stable share throughout the entirety of the period.

4) Voucher Fraud: This type of products brought the most revenue to underground vendors. The majority of the sixteen selected companies in this group are retail stores or chain restaurants, plus a few e-commerce websites. Figure 7 shows that one company greatly surpasses the rest in offering and realized sales: the supermarket chain V1. Their vouchers, which sold on the underground markets had values ranging from $50 to $1000, eventually bringing a cumulative revenue of $203 903 for their vendors. Except for V1, the criminals on the dark web sold vouchers in significantly lower numbers – of no more than a thousand of each of the other companies. However, as the items in this category have on average the highest prices per listing of all the categories, the total revenue is also relatively large.

As can be seen from Figure 7, the highest median price per listing is on vouchers of the e-commerce website A21.2: $285,

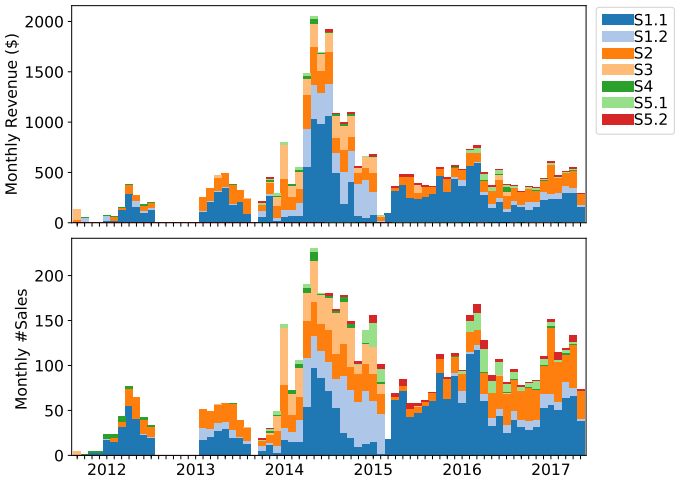


Fig. 6. Number of sales and criminal revenue due to pirated software

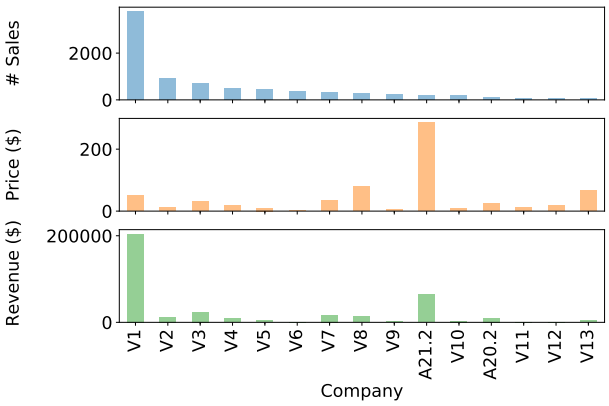


Fig. 7. Revenue, Sales, and Price per Company in "Voucher Fraud" Category

which is also the biggest value across all. This is extremely high value because 70% of their items are $500 gift cards, thus leading to high average and median prices. The second highest price of $80, which is more than three times lower than that of A21.2, is on vouchers for an American retail chain store V8, followed by $67 for gift cards for the online retailer V13. Other widely sold items were gift cards for the restaurant V2, with a comparatively low median price of $12.5; and the coffee store V3, averaging a value of about $30. The lowest price was observed on vouchers for the restaurant chain V6, valued at $2.3.

The variance in pricing of the different vouchers, and the relatively high prices, is due to the relation between their list value and their selling price, as the latter is directly dependent on the former. Unlike the listings in the other categories, which offer similar products, such as accounts for the same service, and therefore are expected to have similar prices, gift cards can vary significantly, as observed earlier. Furthermore, gift cards for large amounts would normally command higher selling prices as well, which could bring about for the bigger prices on average.

Several observations can be made about the distribution of vouchers over time (Figure 8): first, similarly to the accounts category, vouchers for retailers, such as V13, and V11, were

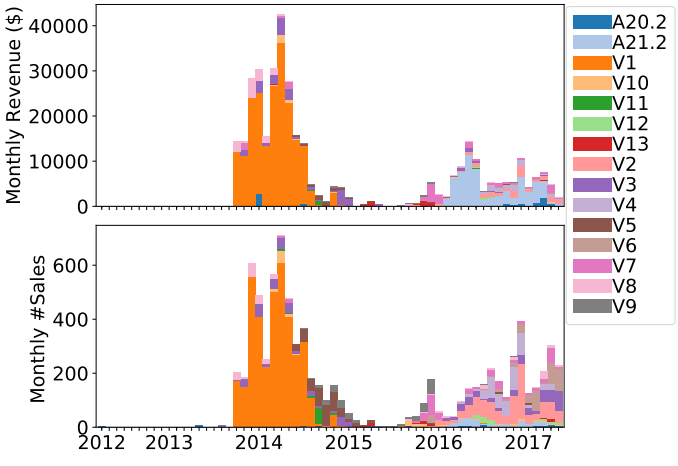


Fig. 8. Number of sales and criminal revenue due to voucher fraud

offered for rather short time intervals. Second, the vouchers for V1, which were the most lucrative item on the examined underground anonymous markets, were distributed during the initial up-rise in the trade, until 2014, and did not turn up again. Finally, a considerable portion of the coupons for restaurants only appeared on Alphabay, after 2016, suggesting a rising interest in this type of vouchers.

B. Underground sales across markets and vendors

To distinguish whether some products were more popular on certain markets or others, we compare the number of sales in the four categories across the examined markets. Figure 9 shows the number of sales per month for the various products. Products from all the categories were offered on the majority of the markets, except for loyalty programs, which appeared on the markets at a later time than the other three. Loyalty programs were primarily traded on Alphabay, with a negligible amount of sales made on Agora and Evolution. The offering in the three categories of accounts, pirated software and vouchers was significant on Alphabay, Silk Road 2 and Evolution, while there was a minimal amount of sales made on Black Market Reloaded, and especially on Pandora.

Sales in the accounts category grossly outnumber transactions in the other three categories. Accounts were also the only ones offered on all eight markets. The offering had its peak in 2014-2015 with the greater part of sales being made on Silk Road 2 and Evolution. After the closure of these markets, and the establishment of Alphabay, there was a considerable amount of accounts traded there, although the volumes do not reach those achieved previously.

The retail of vouchers follows a similar pattern to that of accounts, though on a smaller scale: while the highest amount of accounts sold on one market was over 1200 on Evolution, the most vouchers that were traded on a single market was slightly over 600 on Silk Road 2. The retail of pirated software was consistent over time and markets, with Alphabay reaching the highest number of sales made on a single market, though the number is still slightly behind the highest point of sales made cumulatively in 2014-2015. Overall, it can be seen that items from all categories, except loyalty programs, were consistently offered across the majority of the markets and over the entirety of the examined time period.

Next, we explore the level of vendor competition in the different categories, so as to determine whether there is a variety of sellers supplying the products, or just a few big vendors. The analysis of the distribution of sales across vendors shows that there are a few dominant vendors in each category which have made the greater part of the sales. For the accounts category, in which one of the 183 vendors has been responsible for over half of the sales. The rest is more or less divided between numerous other sellers, with only three or four managing to reach a 5–10% market share.

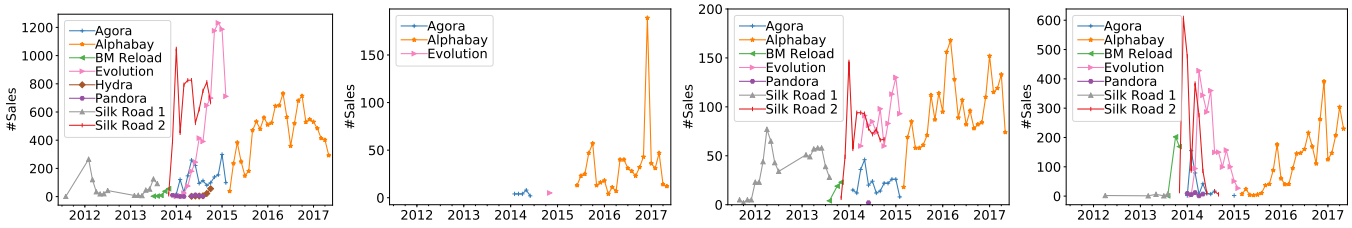
The loyalty programs category sees even less differentiation, though this could also be attributed to the overall much lower number of sellers in this category (18). Nearly 85% of the sales are generated by three vendors, with one making close to 50% of the sales. There have been significantly more vendors active in the pirated software category - 119, though similarly half the market is held by two vendors, who have almost equal market shares. The one category which has more diversity in terms of vendors, is vouchers. Nevertheless, the stronger presence of three – four vendors among the 117 sellers is still evident, though they have reached around 17% market share at most.

In short, a small portion of vendors are responsible for a considerable number of the sales in each category, and consequently - the revenues. Furthermore, the majority of the sellers in each category apparently earn a negligible revenue, hence the earnings from the sale of fraudulent products is probably not the reason they participate in the underground trade.

C. Company Revenue Losses

Revenue loss occurs when a company makes less money than expected due to external and internal factors, in our case, due to the trading taking place in underground markets. Based on the retail prices, the underground price and the number of transactions of the products sold in the underground markets, we estimate revenue losses per company. The revenue loss is calculated per product by comparing the retail price of each item and the price it was sold for on the underground marketplaces. Note that this leads to over-estimating the revenue loss as not all the users that pay for the underground product would have bought the legitimate one. The projected revenue losses incurred by the affected companies from the sale of each item in the underground economy are detailed in Figure 10.

Figure 10 shows that there are differences in revenue losses of even 2 orders magnitudes for companies who suffered from the same fraud type. As expected, the underground price is much lower than the retail price; while in some cases this difference is just a few dollars; some products are offered at prizes 100 times cheaper in the underground market. The largest differences are noted within the loyalty program and the retail stores accounts. The reason for this could be in the way those accounts were valued: according to the listed amount within the account, which significantly raises the estimated revenue losses.



(a) Accounts (b) Loyalty programs (c) Pirated software (d) Vouchers

Fig. 9. Evolution of number of the sales per month across the different marketplaces

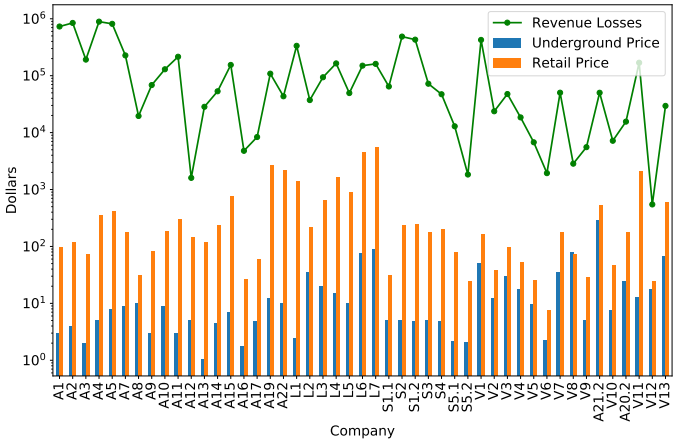


Fig. 10. Company revenue losses, underground and retail product price

V. Explaining Business Fraud Impact

47 companies from different industries and countries had their products being sold in underground marketplaces during the period of analysis. However, one question remains open: what made these companies be the target of this type of fraud? Based on RAT [13], we characterize the value and visibility of these companies to then model the target selection process.

We study the influence of the various company-structural characteristics on business fraud. We create 2 explanatory models: (i) a negative binomial regression model to understand how often a company suffered from business fraud based on its characteristics; and (ii) a linear regression used to examine how these characteristics influence revenue losses.

A. Business fraud and company characteristics

Leveraging the scraped data from the underground markets, we build 2 different dependent variables that capture the impact of business fraud: (i) number of sales of business-fraud-related listings per company; and (ii) total revenue losses per company. The regression models explore the relationship between the impact of business fraud and the company characteristics. As such, the explanatory/independent variables are based on the following company-related factors:

- *Company size*: a continuous variable, estimated as an ensemble of a company's yearly revenue, net income, total number assets, total equity, and number of employees;
- *Popularity*: a continuous variable, based on the domain popularity ranking in Cisco's Umbrella top 1 million list. It is normalized to reflect the variation of the company ranks in the list. In the case a company domain is not featured in the ranking, we assigned 1 million plus one as that company's popularity index (note that only 1 out of the 47 analyzed companies was not present in the top 1 million list);
- *Average underground price*: a continuous variable reflecting the average price a company product was sold for on the underground markets, z-transformed;
- *Area of service*: boolean variable, which receives a value of one when a company operates worldwide, and a zero when it is locally based.

Moreover, we added 2 control variables with the purpose of limiting possible bias resulting from a certain category being too popular in the markets, or the period of time for which an item was offered on the market:

- *Product type*: categorical variable, represented by the four values: account, loyalty program, pirated software, and voucher;
- *Mean lifespan*: continuous variable, reflecting the average number of days a company's item has been for sale.

B. Modelling business fraud

We expect certain degree of correlation among the company characteristics that has to be addressed prior to the modelling. Specifically, several factors reflect one way or another the size of a company, as well as the variables related to visibility, such as reputation or popularity. For instance, a well-established company would be large and have high visibility. Hence, we first check whether the independent variables show high correlation with each other, which could lead to inaccurate coefficient estimations and incorrect interpretations. For this purpose, all independent variables are examined through a Spearman correlation matrix, which is supplied in Figure 11. The correlations with low statistical significance are crossed in the matrix, and we can see the degree of correlations in other cells. It is shown that last six variables with coefficients higher than 0.5 present multicollinearity.

An option for decreasing multicollinearity is by merging the correlated variables into a common dimension through a Principal Component Analysis (PCA). PCA transforms the initial correlated variables into principal components: uncorrelated linear combinations of the original set of variables, weighted by the portion of the variance they explain in the dataset [40].

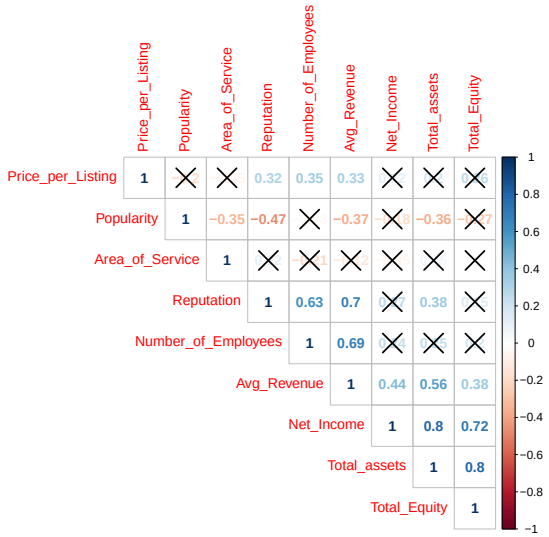


Fig. 11. Spearman Correlation matrix

In order to select the components, we created a scree graph plotting the eigenvalues and the components which suggested the use of three components.

To better fit the data and improve interpretability, we applied a factor rotation. We used the `promax` method which is the commonly used oblique rotation, and looked at the correlations among factors. The rotated component matrix with their 'loadings' in addition to several other quantities of interest is displayed in Table III. The highlighted loadings indicate the variables that are strongly associated with each principal component. We concluded that *net income, total assets and total equity* load on the first component, *number of employees and average revenue* load on the second component, while *reputation* is loading on the third. All three factors exhibit a value higher than 1 for their reported sums of squared loadings ('SS Loadings') and 97% cumulative explained variance; which make them meaningful. Consequently, these three rotated factors are going to be used in the regression analyses.

TABLE III
PCA SUMMARY

Variable	PC_1	PC_2	PC_3
Reputation	0.01	-0.01	**1.01**
Number of Employees	-0.16	**1.07**	-0.01
Avg Revenue	0.27	**0.82**	0.01
Net Income	**1.05**	-0.14	-0.02
Total assets	**0.89**	0.07	0.07
Total Equity	**0.98**	0.05	-0.04
SS Loadings	2.94	1.85	1.01
Proportion var. explained	0.49	0.31	0.17
Cumulative var. explained	0.49	0.80	0.97

Once the multicollinearity issue was solved, we examined the distribution of the dependent variables to determine the most appropriate regression model for the analysis. In our first regression analysis (see Table IV) the *sales* variable is an observed count, following a negative binomial distribution as seen in the descriptive analysis, suggesting a Negative Binomial Regression model would be most suitable. A count variable could also be modelled using a Poisson regression, however that model assumes that the data follows a Poisson distribution and therefore, the mean equals the variance. By observing the standard deviation of 1658.62 of the fraud count variable, it is established that the mean (949.47) is significantly smaller than the variance. This indicates that the data is over-dispersed and a Poisson model would be unsuitable.

Both regression models have the following general structure:

$$ln(d_v) = c_0 + \sum c_i \times v_i + e,$$

where d_v is the dependent variable and v_i are the company-related explanatory variables. The regression coefficients c_i capture the influence of the explanatory variables on the number of sales. Moreover, c_0 is a constant value setting a baseline and finally e an error term. We start by constructing a model which only includes our control variables (model 2) as a baseline to compare against. We construct six additional models (models 3-8) by additionally including the explanatory variables one-by-one to individually demonstrate the effects of them over the number of sales. Finally, we construct the complete model (model 9), which simultaneously includes control variables and other predictors as well.

We observe that 28% of the variance in sales is purely explainable by our control variables. The pseudo-R^2 value of model 2 shows that a significant amount of the variance in sales is explainable by either the lifetime of a listing and its category. Compared to that, 43% of variance in sales is explainable when adding company characteristics to the control variables. This increase in pseudo-R^2 is attributed to the effects of company characteristics. The secondary pseudo-R^2 values relative to model 2 that are presented in Table IV imply that an additional 20% of the variance in fraud transactions is entirely explainable by company characteristics.

Fraud type is represented by three dummy variables in the models where the fourth category (voucher) is the reference variable. The coefficient values associated with each fraud type in model 9 capture the impact on the number of sales. For instance, a change of an item from the *voucher* category to *Loyalty program* category has a $e^{-2.56} = 0.08$ multiplicative decrease in the number of sales, i.e., loyalty program accounts are sold 8% less times than *vouchers*.

The complete model shows that only the Principal component 3 (PC3) appears to have a significant effect on the number of sales. Recall from Table III, PC3 is only correlated with the *reputation* variable. The coefficient value of 0.59 entails that if the remaining explanatory variables are held constant, items that deviate from the mean of PC3 by 1 unit are sold $e^{0.59} = 1.8$ times more. Hence, items belonging to high-reputation companies are sold more often. Similarly, analyzing the impact of popularity separated from reputation (that is excluding PC3; model 4); an increase of 1 unit of popularity will entail $e^{-1.48} = 0.22$ times less sales. That is, increasing the popularity ranking (i.e., less popular) decreases the number of sales. Other company size variables (PC1, PC2, price of the items and Area of service) do not show any significant impact.

Our second model focused on explaining the differences on revenue losses due to business fraud across the different

TABLE IV
GENERALIZED LINEAR REGRESSION MODEL (GLM) FOR NUMBER OF SALES OF ITEMS RELATED TO BUSINESS FRAUD

	Response Variable: Number of sales of products related to business fraud								
	Negative Binomial with Log Link Function								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lifespan		0.001	0.001	0.001	0.002	0.001	0.001	0.004	0.003
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Stolen Accounts		0.73	0.77*	0.43	0.74	0.73	0.78*	0.70	0.32
		(0.46)	(0.46)	(0.45)	(0.47)	(0.47)	(0.46)	(0.43)	(0.47)
Loyalty programs		−1.70***	−1.64***	−1.95***	−1.68***	−1.70***	−1.66***	−2.14***	−2.56***
		(0.58)	(0.61)	(0.58)	(0.59)	(0.58)	(0.59)	(0.56)	(0.60)
Pirated software		0.52	0.57	0.19	0.45	0.54	0.55	0.12	−0.19
		(0.65)	(0.68)	(0.65)	(0.66)	(0.66)	(0.66)	(0.61)	(0.67)
Area_of_Service			0.07						−0.17
			(0.41)						(0.38)
Popularity				−1.48*					−0.93
				(0.87)					(0.90)
Price_per_Listing					0.09				−0.11
					(0.18)				(0.18)
PC1						−0.02			−0.03
						(0.17)			(0.20)
PC2							0.07		−0.21
							(0.17)		(0.21)
PC3								0.53***	0.59***
								(0.16)	(0.20)
Constant	6.86***	6.38***	6.33***	6.69***	6.31***	6.37***	6.34***	6.10***	6.57***
	(0.19)	(0.35)	(0.43)	(0.39)	(0.36)	(0.35)	(0.36)	(0.33)	(0.45)
Pseudo R2	0	0.28	0.28	0.32	0.28	0.28	0.28	0.4	0.43
Pseudo R2 regards to Model2	-	-	0.00058	0.052	0.0026	0.00014	0.0013	0.17	0.2
Observations	47	47	47	47	47	47	47	47	47
Log Likelihood	−365.37	−357.66	−357.65	−356.39	−357.60	−357.66	−357.63	−353.39	−352.28
Akaike Inf. Crit.	732.73	725.32	727.29	724.79	727.20	727.31	727.26	718.79	726.56

Note: *p<0.1; **p<0.05; ***p<0.01

companies. We use a log-linear regression model to capture the impact of a company's characteristics on revenue losses. Table V presents a summary of the different models. 18% of the variance in revenue losses is purely explainable by the control variables, and an additional 29% is merely explainable by company characteristics.

Pirated software and *Loyalty program* items have a positive statistically significant impact in the total revenue losses. For example, if all else were held constant, a change of an item from *voucher* to *pirated software* is correlated with a $e^{2.12} = 8.33$ multiplicative increase in the total revenue losses. Thus, *pirated software* and *Loyalty program* create 733% and 426% ($e^{1.66} = 5.26$) higher revenue losses compared to *voucher* fraud. Among the company characteristics, *PC3* and *Area of service* have a significantly effect on revenue losses. Items that deviate from the mean of PC3 variable by 1 unit have $e^{0.62} = 1.86$ times more revenue losses, i.e., high-reputation companies lose more revenue because of the business fraud happening on the underground market; which is also related with their products being highly exposed to fraud as shown in our first regression. Surprisingly, local companies (Area of service = 0) experience more financial losses relative to companies with global presence. Specifically, companies operating locally exhibits $e^{1.40} = 4.06$ times more revenue losses than operating worldwide. Other explanatory variables do not demonstrate significant impact on revenue losses.

VI. DISCUSSION

a) Recommendations: Business fraud is evolving with new types of products being sold in underground markets ranging from video streaming accounts to gift cards and air miles. While this type of products are rarely associated with high fraud risk; companies are facing considerable revenue losses. The identification and categorization of the different types of business fraud can aid companies in determining more easily which of their offered products could be exploited: gift cards, accounts, etc.; and take the appropriate contingency measures. Monitoring underground markets will provide valuable threat intelligence that could lead to the creation of more robust service fraud detection and prevention methods, as well as more strict consumer protection regulations, similar to those existing in the mature financial sector.

Our statistical models indicated that established companies are at a higher risk of being targeted, while the relatively small businesses may suffer more losses. Larger companies may have implemented more appropriate fraud detection and mitigation measures, or are better at absorbing the inflicted revenue losses. The results of our models could support decision making and risk assessment evaluations. The majority of existing cyber risk management frameworks make use of expected losses and risk probability estimations to quantify the potential risks a company is exposed to. The company size metric, as well as the reputation and location characteristics, after further refinement, could be applied in such risk assessments, facilitating the cyber risk management process.

b) Limitations: Despite analyzing all the listings of 8 major marketplaces for 8 years, we were only able to identify 47 companies who suffered from business fraud. Yet, the estimated figures presented with regard to the impact of business fraud must be understood within the assumptions of the analysis due to the following reasons. First, we chose the lowest retail price for every product offered in the anonymous marketplaces. Second, while the dataset has been extensively by other researchers [9], [10], it is scraped data and as such fundamentally not ground truth. Finally, the number of transactions was based on the number of feedbacks which represent a lower estimation of the transactions.

Additional business fraud might have been facilitated via

TABLE V
LOG-LINEAR MODEL FOR TOTAL REVENUE LOSSES PER COMPANY

	Response Variable: total revenue losses								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Lifespan		0.01	0.01	0.01	0.01	0.005	0.01	0.01	0.01*
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Stolen accounts		1.00	1.21	1.01	0.74	1.16	0.93	0.89	1.05
		(0.75)	(0.73)	(0.77)	(0.76)	(0.76)	(0.76)	(0.75)	(0.77)
Loyalty Programs		1.81*	2.47**	1.83*	1.55	1.77*	1.74*	1.48	1.66*
		(0.95)	(0.96)	(0.97)	(0.95)	(0.95)	(0.96)	(0.97)	(0.98)
Pirated Software		1.68	2.33**	1.71	1.39	1.81*	1.59	1.58	2.12*
		(1.07)	(1.07)	(1.10)	(1.07)	(1.07)	(1.08)	(1.06)	(1.09)
Area_of_Service			1.40**						1.40**
			(0.64)						(0.62)
Popularity				0.23					0.77
				(1.46)					(1.47)
Price_per_Listing					0.48				0.41
					(0.29)				(0.29)
PC1						−0.32			−0.47
						(0.28)			(0.33)
PC2							−0.20		−0.13
							(0.29)		(0.34)
PC3								0.39	0.62*
								(0.28)	(0.32)
Constant	10.65***	9.24***	8.42***	9.19***	9.06***	9.23***	9.32***	9.23***	8.12***
	(0.29)	(0.58)	(0.67)	(0.67)	(0.58)	(0.58)	(0.60)	(0.57)	(0.73)
Pseudo R2	0	0.18	0.27	0.18	0.23	0.21	0.19	0.22	0.42
Pseudo R2 regards to Model2	-	-	0.11	0.00064	0.061	0.031	0.012	0.044	0.29
Log Likelihood	-98.17	-93.5	-90.9	-93.49	-92.04	-92.77	-93.22	-92.47	-85.79
Observations	47	47	47	47	47	47	47	47	47
Akaike Inf. Crit.	200.33	199.00	195.79	200.97	198.08	199.54	200.45	198.93	195.58

Note: *p<0.1; **p<0.05; ***p<0.01

underground forums. To understand the coverage and representativeness of our data, we also looked at CrimeBB [41]. However, since most of the trading is performed privately in 'private messages', it is impossible to quantify the number of sales in these forums.

In addition, while the feedback left by the buyers indicated that the bought products were satisfactory and enabled them to commit fraud, we do not have hard evidence that the business fraud was committed in all cases. Given the nature of some of the products (accounts, vouchers, loyalty programs) that could be reused or that have zero marginal production costs, it is unclear what the actual losses to the companies were. Thus, revenue losses estimates must be also taken with caution.

c) Ethics: All data used in our study was gathered from publicly available sources as anyone could access the underground markets when they were online. Given the anonymous nature of these marketplaces, the data did not contain any type of Personal Identifiable Information (PII). This study was exempted by our Institutional Review Board (IRB) since the datasets were publicly available. With regards to the affected companies, their names have been anonymized.

VII. CONCLUSIONS

Previous research has mostly focused on consumers as victims of fraud; but the reality is that businesses are also frequently victims of fraud. These types of fraud are diverse both in methods and perpetrators, from scams carried out by employees to fraudulent returns from customers to identity theft by outsiders. Cybecriminals have recently expanded their business from committing the fraud themselves to act as facilitators. In this paper, we have focused on quantifying this type of fraud by which criminals offer illegal products and services to consumers via anonymous online markets.

By analyzing scraped data of 8 major underground marketplaces, we identified 47 companies whose products were sold illegally. Fraudulent consumer products that facilitate business fraud were consistently traded on the eight examined underground markets. Our analysis showed that criminals made a relatively low revenue out of this selling; while businesses suffer major revenue losses ranging from a few hundred thousand up to a million dollars. In total, we estimated that the total revenue losses for these businesses reached about 7.5 million dollars for the whole period of analysis.

We classified these products into 4 different categories: stolen accounts, loyalty programs, vouchers and pirated software. Selling pirated software together with loyalty programs were the most lucrative products for criminals, while accounts make up the greatest part of all transactions which took place. Most of these products had very short lifetime, being the median offering time below 30 days, though some products were observed for longer periods of more than a year.

In addition, we have presented 2 statistical models to measure the relationship between structural characteristics of the affected companies and the amount of products being sold in the underground markets. Our results showed that criminals are selling more products from companies with higher visibility and reputation that from companies not that well-known. Smaller companies operating regionally are the ones who suffer the highest revenue losses. Surprisingly, the difference between the mean price in its category and the price the product sold in the underground market does not seem to affect the number of transactions.

References

[1] K. Bissell, R. LaSalle, and P. D. Cin, "The cost of cybercrime," Accenture and Ponemon Institute, Tech. Rep., Mar. 2019. [Online]. Available: https://www.accenture.com/_acnmedia/PDF-96/Accenture-2019-Cost-of-Cybercrime-Study-Final.pdf#zoom=50

[2] McAfee and CSIS, "Net losses: Estimating the global cost of cybercrime," Tech. Rep., Jun. 2014. [Online]. Available: https://csis-website-prod.s3.amazonaws.com/s3fs-public/legacy_files/files/attachments/140609_rp_economic_impact_cybercrime_report.pdf

[3] P. Dreyer, T. Jones, K. Klima, J. Oberholtzer, A. Strong, J. Welburn, and Z. Winkelman, "Estimating the global cost of cyber risk: Methodology and examples," Tech. Rep., 2018. [Online]. Available: https://www.rand.org/pubs/research_reports/RR2299.html

[4] C. Gañán, M. Ciere, and M. Van Eeten, "Beyond the pretty penny: The economic impact of cybercrime," in *Proceedings of the New Security Paradigms Workshop (NSPW '17)*, Santa Cruz, CA, USA, Oct 2017, p. 35–45. [Online]. Available: https://doi.org/10.1145/3171533.3171535

[5] R. Anderson, C. Barton, R. Bölme, R. Clayton, C. Gañán, T. Grasso, M. Levi, T. Moore, and M. Vasek, "Measuring the changing cost of cybercrime," in *Proceedings of Workshop on Economics of Information Security (WEIS '19)*, 2019. [Online]. Available: https://weis2019.econinfosec.org/wp-content/uploads/sites/6/2019/05/WEIS_2019_paper_25.pdf

[6] L. Zhang-Kennedy, H. Assal, J. Rocheleau, R. Mohamed, K. Baig, and S. Chiasson, "The aftermath of a crypto-ransomware attack at a large academic institution," in *Proceedings of 27th USENIX Security Symposium (USENIX Security 18)*, Baltimore, MD, USA, 2018, pp. 1061–1078.

[7] D. Palmer. (2018, Oct.) This is how much the WannaCry ransomware attack cost the NHS. [Online]. Available: https://www.zdnet.com/article/this-is-how-much-the-wannacry-ransomware-attack-cost-the-nhs/

[8] L. Mathews, "NotPetya ransomware attack cost shipping giant Maersk over $200 million," *Forbes*, Aug. 2017. [Online]. Available: https://www.forbes.com/sites/leemathews/2017/08/16/notpetya-ransomware-attack-cost-shipping-giant-maersk-over-200-million/#486580244f9a

[9] K. Soska and N. Christin, "Measuring the longitudinal evolution of the online anonymous marketplace ecosystem," in *Proceedings of 24th USENIX Security Symposium (USENIX Security '15)*, Washington, D.C., Aug. 2015, pp. 33–48. [Online]. Available: https://www.usenix.org/conference/usenixsecurity15/technical-sessions/presentation/soska

[10] R. Van Wegberg, S. Tajalizadehkhoob, K. Soska, U. Akyazi, C. Gañán, B. Klievink, N. Christin, and M. Van Eeten, "Plug and prey? Measuring the commoditization of cybercrime via online anonymous markets," in *Proceedings of 27th USENIX Security Symposium (USENIX Security 18)*, Baltimore, MD, USA, Aug. 2018, pp. 1009–1026.

[11] R. van Wegberg, F. Miedema, U. Akyazi, A. Noroozian, B. Klievink, and M. van Eeten, "Go see a specialist? Predicting cybercrime sales on online anonymous markets from vendor and product characteristics," in *Proceedings of The Web Conference 2020*, Taipei, Taiwan, Apr. 2020, p. 816–826.

[12] A. Hutchings, "Leaving on a jet plane: the trade in fraudulently obtained airline tickets," *Crime, Law and Social Change*, vol. 70, no. 4, pp. 461–487, May 2018.

[13] M. Yar, "The novelty of 'cybercrime': An assessment in light of routine activity theory," *European Journal of Criminology*, vol. 2, no. 4, pp. 407–427, Oct. 2005.

[14] D. Thompson, "1997 computer crime and security survey," *Information Management & Computer Security*, vol. 6, no. 2, pp. 78–101, May 1998.

[15] K. Fanning, "Minimizing the cost of malware," *Journal of Corporate Accounting & Finance*, vol. 26, no. 3, pp. 7–14, 3 2015. [Online]. Available: https://doi.org/10.1002/jcaf.22029

[16] The Council of Economic Advisers, "The cost of malicious cyber activity to the U.S. economy," Executive Office of the President of USA, Tech. Rep., Feb. 2018. [Online]. Available: https://www.whitehouse.gov/articles/cea-report-cost-malicious-cyber-activity-u-s-economy

[17] L. Paoli, E. Van Hellemont, C. Verstraete, J. Visschers, R. De Wolf, M. Martens, L. De Marez, P. Verdegem, E. Teerlinck, P. Chen, C. Huygens, T. De Cnudde, V. Rijmen, M.-C. Janssens, and T. Marquenie, "Belgian cost of cybercrime: Measuring cost and impact of cybercrime in Belgium," Belgian Science Policy Office, Tech. Rep., 2018. [Online]. Available: http://www.belspo.be/belspo/brain-be/projects/FinalReports/BCC_Final%20Report.pdf

[18] C. Herley and D. Florêncio, "Nobody sells gold for the price of silver: Dishonesty, uncertainty and the underground economy," in *The 9th Workshop on the Economics of Information Security (WEIS '10)*. Cambridge, MA, USA: Springer, Jun. 2010, pp. 33–53.

[19] R. Anderson, C. Barton, R. Böhme, R. Clayton, M. J. G. Van Eeten, M. Levi, T. Moore, and S. Savage, "Measuring the cost of cybercrime," in *Proceedings of The 11th Workshop on the Economics of Information Security (WEIS '12)*, Berlin, Germany, Jun. 2012, pp. 265–300.

[20] M. Lagazio, N. Sherif, and M. Cushman, "A multi-level approach to understanding the impact of cyber crime on the financial sector," *Computers & Security*, vol. 45, pp. 58–74, Sep. 2014.

[21] A. Noroozian, J. Koenders, E. Van Veldhuizen, C. Gañán, S. Alrwais, D. McCoy, and M. Van Eeten, "Platforms in everything: analyzing ground-truth data on the anatomy and economics of bullet-proof hosting," in *Proceedings of the 28th USENIX Conference on Security Symposium*, 2019, pp. 1341–1356.

[22] Verizon Enterprise, "2015 data breach investigation report," Tech. Rep., 2015. [Online]. Available: https://enterprise.verizon.com/resources/reports/2015/data-breach-investigation-report_2015.pdf

[23] M. Riek, R. Boehme, M. Ciere, C. Gañán, and M. van Eeten, "Estimating the costs of consumer-facing cybercrime: A tailored instrument and representative data for six eu countries." in *Workshop of Economics of Information Security (WEIS '16)*, 2016.

[24] M. Riek and R. Böhme, "The costs of consumer-facing cybercrime: an empirical exploration of measurement issues and estimates," *Journal of Cybersecurity*, vol. 4, no. 1, Oct. 2018. [Online]. Available: https://doi.org/10.1093/cybsec/tyy004

[25] J. Lewis, "The economic impact of cybercrime-no slowing down," Tech. Rep., Feb. 2018. [Online]. Available: https://csis-website-prod.s3.amazonaws.com/s3fs-public/publication/economic-impact-cybercrime.pdf

[26] Detica, "The cost of cyber crime," Tech. Rep., 2011. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/60943/the-cost-of-cyber-crime-full-report.pdf

[27] K. Thomas, D. Huang, D. Wang, E. Bursztein, C. Grier, T. J. Holt, C. Kruegel, D. McCoy, S. Savage, and G. Vigna, "Framing dependencies introduced by underground commoditization," in *Workshop on Economics of Information Security (WEIS '15)*, Delft, The Netherlands, Jun. 2015.

[28] C. H. Malin, T. Gudaitis, T. J. Holt, and M. Kilger, "Social dynamics of deception," in *Deception in the Digital Age*. Elsevier, 2017, pp. 125–148. [Online]. Available: https://doi.org/10.1016/B978-0-12-411630-6.00004-9

[29] R. Thomas and J. Martin, "The underground economy: priceless," *login: The Usenix Magazine*, vol. 31, no. 6, Dec. 2006. [Online]. Available: https://www.usenix.org/system/files/login/articles/822/cymru.pdf

[30] M. Motoyama, D. McCoy, K. Levchenko, S. Savage, and G. M. Voelker, "An analysis of underground forums," in *Proceedings of the ACM SIGCOMM conference on Internet measurement conference (IMC '11)*, Berlin, Germany, Nov. 2011, pp. 71–80.

[31] T. J. Holt, O. Smirnova, and A. Hutchings, "Examining signals of trust in criminal markets online," *Journal of Cybersecurity*, vol. 2, no. 2, pp. 137–145, 2016.

[32] J. Broséus, D. Rhumorbarbe, C. Mireault, V. Ouellette, F. Crispino, and D. Décary-Hétu, "Studying illicit drug trafficking on darknet markets: Structure and organisation from a Canadian perspective," *Forensic Science International*, vol. 264, pp. 7–14, 7 2016. [Online]. Available: https://doi.org/10.1016/j.forsciint.2016.02.045

[33] M. Yip, C. Webber, and N. Shadbolt, "Trust among cybercriminals? Carding forums, uncertainty and implications for policing," *Policing and Society*, vol. 23, no. 4, pp. 516–539, Apr. 2013. [Online]. Available: https://doi.org/10.1080/10439463.2013.780227

[34] N. Christin, "Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace," in *Proceedings of the 22nd International Conference on World Wide Web*, May 2013, pp. 213–224. [Online]. Available: https://doi.org/10.1145/2488388.2488408

[35] M. J. Barratt, "Silk road: Ebay for drugs," *Addiction*, vol. 107, no. 3, p. 683, Feb. 2012. [Online]. Available: https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1360-0443.2011.03709.x

[36] J. Buxton and T. Bingham, "The rise and challenge of dark net drug markets," *Policy Brief*, vol. 7, pp. 1–24, 2015.

[37] Europol, "Internet organised crime threat assessment," *Network Security*, vol. 2018, no. 10, p. 4, Oct. 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1353485818300965

[38] Impact - Cyber Trust. (2020) Information marketplace for policy and analysis of cyber-risk & trust. [Online]. Available: https://www.impactcybertrust.org/

[39] Federal Bureau of Investigation. Business fraud. [Online]. Available: https://www.fbi.gov/scams-and-safety/common-scams-and-crimes/business-fraud

[40] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433–459, 2010. [Online]. Available: https://doi.org/10.1002/wics.101

[41] S. Pastrana, D. R. Thomas, A. Hutchings, and R. Clayton, "Crimebb: Enabling cybercrime research on underground forums at scale (WWW '18)," in *Proceedings of the World Wide Web Conference*, Lyon, France, Apr. 2018, pp. 1845–1854.