# FRAUD CHARGEBACK DETECTION FOR STARTUPS BASED ON USER ONLINE PAYMENTS PERFORMANCE

by

Danyil Yaremchuk

A thesis submitted in partial fulfillment of the requirements for the degree of

MA in Business and Financial Economics

Kyiv School of Economics

2021

Thesis Supervisor: ______________________________ Professor Serhiy Gvozdiov

Approved by
Head of the KSE Defense Committee, Professor [Type surname, name]

Date ______________________________________

## TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF TERMS AND ABBREVIATIONS

**Bank acquirer** is a bank (financial institution) that opens an account for the seller to receive transactions from buyers and participates in all financial transactions related to the activities of the business

**Cardholder** owner of the card, client of the bank that issued card

**Card Network** is an international payment system. Example: Visa, MasterCard, Discover, etc

**Cascade** redirecting payments that did not go through one provider to another

**Chargeback** is a procedure for protesting a transaction by the issuing bank (from the client's side), in which the payment amount is debited from the recipient and returned to the payer

**FFF** Family, Friends & Fools

**Fraud** is a card fraud aimed at illegal use of money from her account

**GII** Global Innovation Index

**Issuing bank** is the bank that issued the buyer's card

**NGO** Non-Governmental Organization

**PSP** Payment Service Provider, a provider that accepts payments

**Startup** it is a company or temporary organization created to find a repetitive and scalable business model

**VPN** Virtual Private Network

**3D secure** method of verification when making a payment. Shifts responsibility for fraud to the bank that issued the buyer's card. For example, SMS code.

# CHAPTER 1. INTRODUCTION

Many startups According to the “Startup Ecosystem Rankings” report, made by StartupBlink company, Ukraine was in 29th position in the top ecosystems for startups worldwide. In this study, Kyiv is at 32nd place on the Global Rank of cities, which makes it a European startup hub. Also, Ukraine is at 40th place on the number of startups in the country's top tier made by “Startup Ranking”.

Thus, the startup industry in Ukraine is successful and has had rapid growth in recent years. That makes research about Ukrainian startups interesting, useful and an actual topic to study.

Many startups develop applications, websites and are selling their products online. For businesses to be able to accept card payment or another type of online payments they need to plug into payments providers or develop their own infrastructure. Wrong decisions in this field can cause revenue loss of up to 30%.

Another problem that startups face after they are able to accept online payments is fraud. Fraud is, unfortunately, a common social phenomenon that highly affects the online payments industry. Cardholders whose card information was stolen and payment was performed usually issue a chargeback in order to get their money back. On average businesses are losing 5% of their revenues due to fraud. Cybercrimes are hard to detect and punish because of their international origin and easiness to confuse the crime traces.

In addition to a revenue loss from fraud, businesses with high chargeback rates are included in the card networks (like VISA or MasterCard) monitoring program. If a company has been in the monitoring program for more than 3 months, it can be blocked from accepting online payments in the future.

In this research, we are going to analyze one of Ukrainian startups' payment performance from September 2020 to September 2021 to find patterns in online payment user performance that can signal that transaction is fraud. Based on these factors startups can develop their own anti-fraud system based on risk rules or machine learning that decreases the amount of fraud and chargeback on their products.

# CHAPTER 2. INDUSTRY OVERVIEW AND RELATED STUDIES

## 2.1 Startup Industry Overview

The startup industry in Ukraine has been growing rapidly in the last few years. In the 2020 "Startup Ecosystem Rankings" report, made by StartupBlink company which ranks the startup ecosystems of 1000 cities and 100 countries, Kyiv became a European hub, number 8 at European sites ranking (see table 2.1) and Ukraine is in 29th position (see table 2.2) at the world's startup ecosystem ranking.

Table 2.1 Ukrainian cities ranking

| National Rank | City | Global Rank | Rank Change (from 2019) | Total Score |
|---|---|---|---|---|
| 1 | Kyiv | 32 | +2 | 9.712 |
| 2 | Lviv | 354 | -55 | 0.452 |
| 3 | Odessa | 356 | -121 | 0.450 |
| 4 | Kharkiv | 441 | -6 | 0.318 |
| 5 | Ternopil | 724 | -46 | 0.122 |

The report calls the Ukrainian startup ecosystem "truly inspiring" because despite our country is going through economic problems, it still manages to scale global technologies. Authors of the report mansion the high quality of Ukrainian developers which are in demand by many foreign companies as one of the main factors of the success. Another factor is the mentality of Ukrainians, specifically that notwithstanding the low cost of living and high salaries for developers they still give up this easy cash and create their own startups. They say if the trend continues, Ukraine can have bigger success and even take a leadership role in the global startup ecosystem.

Table 2.2 World countries ranking

| Rank | Country | Rank Change (from 2019) | Quantity Score | Quality Score | Business Score | Total Score |
|---|---|---|---|---|---|---|
| 28 | Austria | — | 1.28 | 0.59 | 3.21 | 5.080 |
| 29 | Ukraine | +2 | 0.84 | 2.18 | 2.04 | 5.057 |
| 30 | Taiwan | new | 0.91 | 1.51 | 2.63 | 5.044 |

However, the report is also pointing out that Ukrainian government is not able to support the startup ecosystem while benefiting from its success. The authors state that Ukrainian public sector should finance the startup infrastructure in order not to lose the entrepreneurs because the temptation to leave Ukraine and immigrate to European countries, for example Poland, is still strong especially when a visa to Europe is no longer needed. Thus, the brain drain is the main risk for Ukrainian startup industry is facing.

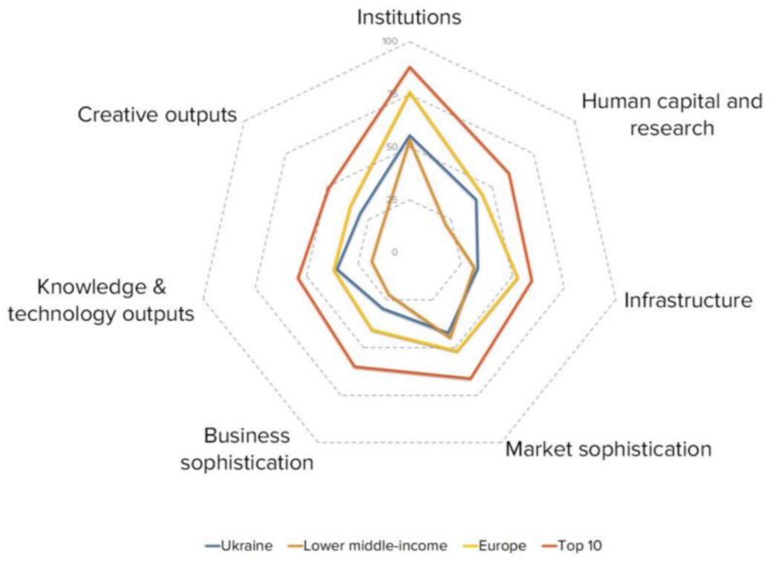
During the last five years, the startup sector increased more than 10 times its capitalization from $ 39 million in 2014 to $ 509 million in 2019. In Ukraine has been more than 146000 patents have been registered since 2007.

The most famous examples of Ukrainian startups which have a massive member base are “Ahrefs”, “MacPaw”, “PetCube”, “Netpeak”, “TemplateMonster”, “Ajax Systems”, “People.ai”, “Reface”, “Reastrem”, “Preply” and “Grammarly”, “Bitfury”, “GitLab” have been valued at more than $1 billion each. Most of them are listed in 2019 “Ukrainian Startups Wall of Flame”, which means that their growth is from 80% to 100%

YoY, $ 1 million annual revenue or $5 million total funding, more than 10 employees in Ukraine.

In the 2020 Global Innovation Index report, Ukraine is in 45th place, which is a bit worse result comparing to previous years, Ukraine was in 42nd place. The strengths and weaknesses of Ukraine in the Global Innovation Index you can see at Figure 2.1. We can see that Ukraine has high scores in pillars such as pillars Institutions, Human capital and research, Infrastructure, Business sophistication, Knowledge and technology, and Creative outputs compared to the lower-middle-income group, which is above the average. In Market sophistication, Ukraine is below the average.

Figure 2.1 Ukraine's scores in the seven GII pillars by Global Innovation Index 2020



Source: https://www.wipo.int/edocs/pubdocs/en/wipo_pub_gii_2020/ua.pdf

## 2.2 Online Payment System Industry Overview

The coronavirus pandemic has accelerated the development of the online payments industry by two to three years. People have a trend to buy goods and services online, where they pay with bank cards or other online payment methods.

In 2020, the volume of the digital payments market is $ 79.3 billion and is expected to grow to $ 154.1 billion in 2025. The main reasons are the spread of the Internet, smartphones and online commerce among the population. This stimulates the development of new solutions in the areas of processing, payment protection and anti-fraud systems.

Three trends characterize the payments market now:

Debit and credit cards are replacing cash and checks offline. This is especially accelerated in connection with contactless payment technologies. Mobile contactless payments estimated at $ 131.36 billion in 2020;

Internet retailing is growing. In 2020, it accounted for 14.4% of the total retail trade in the United States;

Increase in the number of fraudulent activities in online payments.

The main factor holding back the online payments industry is the lack of global standards for international payments. The development of international trade stimulates an increase in the number of international online payments. But different countries have different storage rules and different payment requirements. This increases the accumulation of inefficiencies in the industry.

An opportunity for the development of the sphere of online payments is the introduction of the Open-Banking API. It will allow you to have secure access to customer information for participants in the process of online transactions. It will also simplify the accounting statement for companies and lower fees for processing online payments.

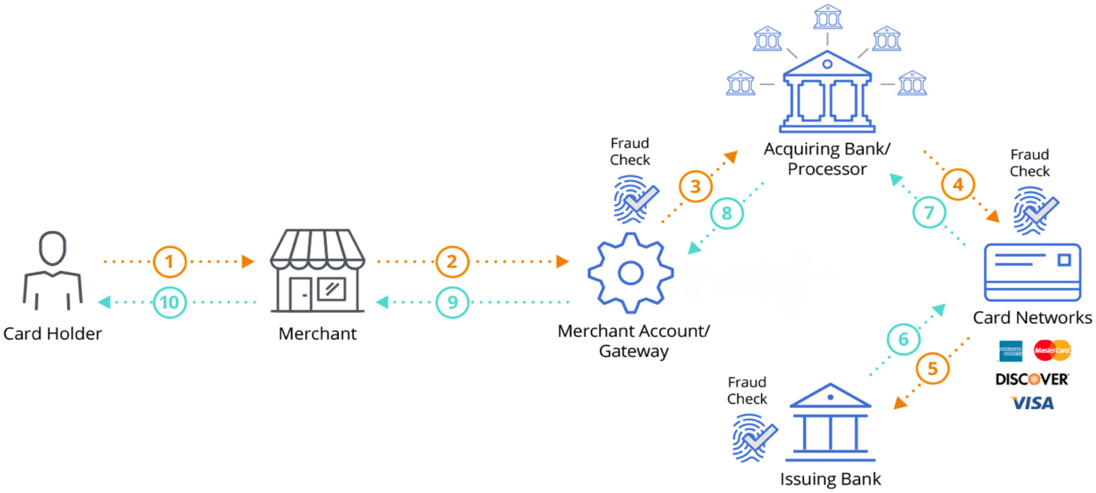
One of the biggest challenges for the industry is the increase in cyberattacks. Market regulators are increasingly paying attention to compliance with confidentiality and information leakage requirements. In addition to fraud, it is also money laundering, DDoS attacks, hacking of payment systems in order to obtain card data and data on cardholders, and more. According to the Central Statistics Office (CSO), losses from cyberattacks will amount to 6 trillion US dollars, which will also be selected for the online payments sector as a whole. AFP Payments Fraud and Control Survey Report of 2018 reports that over 86% of organizations have encountered cyber attacks. It follows from this that the growth of cyberattacks and fraud can slow down the development of services in the field of online payments.

There are several types of transactions: first, token payments and moto payments. The first payments occur when the user enters his card details on the form on the site. Further, money write-offs can be carried out using the generated token, which allows you to make payments without entering bank card details.

Figure 2.2 shows a simplified algorithm of how Payment Processing takes place and the stages at which anti-fraud checks pass.

The cardholder wants to purchase some goods from the merchant. He/she selects goods and fills out a payment form. Next, the payment is processed at the Gateway, which has its anti-fraud system. Then the payment goes to the bank's acquirer, sometimes they also have their anti-fraud system. After the payment is processed in card networks such as VISA, MasterCard, etc., there is also a check for fraud. The payment goes to the issuing bank, which announces the decision to process the payment based on its corporate rules and anti-fraud system. Then, along the chain, the payment information is returned back to the cardholder and the Merchant.

Figure 2.2 Marketplace payment process



Source: https://home.bluesnap.com/payment-processing/

The main types of fraud in online payments are:

- Card check. Fraudsters try to check the validity of the cards and the availability of money on them in order to resell them on the black market in the future.
- Product Specific Fraud (Banking, etc.). Fraudsters try to sell money from stolen cards by obtaining real goods or services.
- Marketing Fraud. Often occurs when working with CPA grids.
- Arbitragers Fraud. They are trying to dilute low-quality traffic with payments from stolen or their own cards, which will soon be used for black-and-white paper.
- Account Takeover. Account theft is becoming one of the most “popular” types of fraud.

- Friendly and Family Fraud. Friendly-Fraud is done by ordinary users who are dissatisfied, disappointed, or just know how the system works and want to get their money back when the service has already been provided to them. The transactions of such users are completely normal and difficult to identify in any way. A high sum of all purchases, followed by a sharp drop in activity on the product.

## 2.3 Related studies

Analytical material: MAIN TRENDS OF STARTUP DEVELOPMENT IN UKRAINE PROBLEMS, OBSTACLES AND OPPORTUNITIES

By: Babiachok R., Kulchytsky I.

This analytical material was prepared within the project № 51321 "Strengthening the impact of the public on European integration processes in the field of science, technology and innovation Development of Ukraine". Babiachok R. is entrepreneur-innovator, and Kulchytsky I. is an expert of the NGO "Agency for European Innovation".

The most frequently startups are classified through 5 stages of development:

1) Seed stage

This is the stage where the ideas about product, service, technology are born. In this stage the planning of developing ideas and finding are discussed as well as a team is formed. The seed stage startup already needs an investment. Usually it is financed by 3F (Family, Friends and Fools). In this stage startups participate in accelerator programs on startup incubators. At this stage it is the most difficult to get investment because there is no evidence about the startup's future efficiency.

2) Startup stage

At this stage, a startup already should have a business model, strategy for promoting the project on the market, a well-coordinated team with clearly defined functions of each member, a legal entity, the first steps in advertising and finding investors.

3) Growth stage

This is the best stage to attract investor's financing. Startup is at final formation of the product to bring it to the perfect condition. The company should already exist, take market share and have income. Investments are still needed because the company does not always have enough profit to grow or have no profit at all.

4) Expansion stage

At this stage a company has the final product and is making profit. Marketing strategy is crucial and needs an expansion. This is the stage when founders could think about selling their company.

5) Exit stage

At this stage founders of the company begin to issue shares or sell it. Startup already is a massive company, has a large market share and is very profitable.

The analytical material highlights main resources for startups: own funds, loans, crowdfunding, participation in competitions, investor funds, venture financing, business incubators and acceleration programs.

Conference Paper: Identifying online credit card fraud using Artificial Immune Systems

By: A. Brabazon, J. Cahill, P. Keenan, D. Walsh

This article discusses approaches to detecting credit card fraud. It describes the challenges of detecting fraudsters, because they are constantly adapting their strategies. Based on data of the payments, the article describes the implemented three artificial immune systems (AIS) algorithms capable of detecting fraud. It indicated patterns of behavior for training the model.

# CHAPTER 3. METHODOLOGY

## 3.1 Hypothesis

In this research following twelve hypotheses are tested:

1. Orders with higher amounts are more likely to be made by fraudsters and chargebacked by cardholders.

2. Transactions with a higher sequence number are more likely to be fraudulent and be chargebacked by cardholders.

3. Transactions that have a higher cascade number (had a decline on previous psp) are more likely to be made by frauders.

4. Users that spend less time on a payment form filin card data are more likely to be frauders.

5. Transactions that have been verified by 3D secure protocol are less likely to be made by frauders.

6. Transactions that are made from a different country than the country in which the card was issued are more likely to be made by fraudsters.

7. Transactions from a cardbrand MasterCard are less likely to be fraud than a VISA cardbrand.

8. Transactions from a cardbrand MasterCard are less likely to be fraud than another cardbrands (not VISA).

9. Bebit cards have fewer fraudulent transactions than credit cards.

10. Debit cards have fewer fraudulent transactions than prepaid cards.

11. Debit cards have fewer fraudulent transactions than other types of cards (not debit or prepaid).

12. Users who enter the name of the cardholder into the form are less likely to be frauders.

### 3.2 Model and variables

In the research, we are testing twelve hypotheses mentioned above. Binary response model is appropriate because we are dealing with a binary variable: chargeback flag — "1" if a transaction had got a chargeback or "0" if a transaction had not got a chargeback. The logit model is well suited for testing hypotheses mentioned above.

Models which has been processed in this research is:

$$CHB = ln(A) + ln(ON) + CN + ln(TF) + D + CM + CBV + CBO +$$

$$+CTD + CTP + CTO + NW$$

where CHB — flag if order had a chargeback "Yes" or "No";

A — amount of order in cents USD;

ON — number of customers "first" order;

CN — order's number of cascade psp;

TF — time that user spent on payment form in seconds;

D — dummy variable, "1" if transaction is 3D secure and "0" if it is 2D;

CM — dummy variable, “1” if there is no country simmatch for transaction (IP location of user and card issuer country are the same) and “0” if there is country mismatch;

CBV — dummy variable, “1” if cardbrand is VISA and “0” if cardbrand is MasterCard;

CBO — dummy variable, “1” if cardbrand is other (not VISA or Mastercard) and “0” if cardbrand is MasterCard;

CTD — dummy variable, “1” if card type is debit and “0” if cardbrand is credit;

CTP — dummy variable, “1” if card type is prepaid and “0” if cardbrand is credit;

CTO — dummy variable, “1” if card type is other (not credit, debit or prepaid) and “0” if cardbrand is credit;

NW — dummy variable, “1” if user has entered cardholder name in the payment form and “0” if if user has not entered cardholder name in the payment form;

### 3.3 Logit model

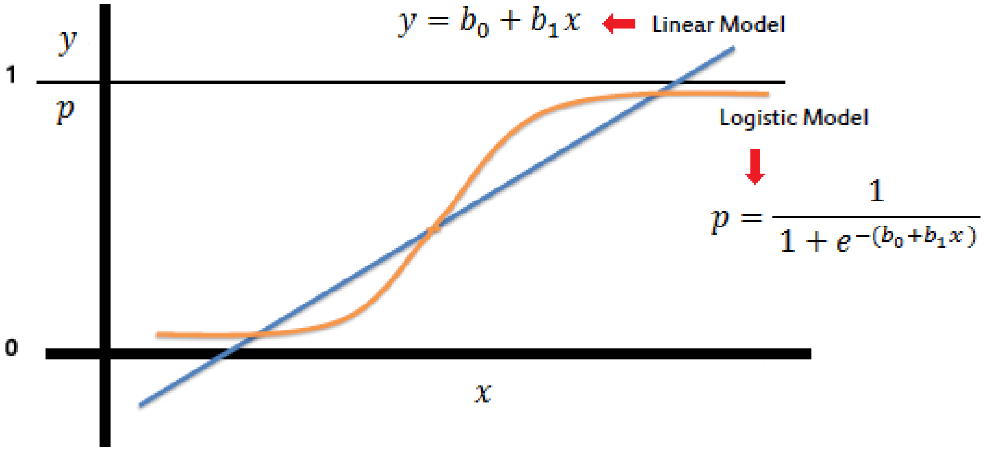
The logit model is an alternative to the linear probability model that helps to get rid of the LPM disadvantages like constant partial effects and probabilities less than 0 and greater than 1.

In order to avoid LPM limitations consider a function 0<G(z)<1 for all real numbers z, which:

$$P(y = 1|X) = P(\beta 0 + \beta 1X2 + \cdots + \beta kXk + u) = G(z)$$

G=exp(z)/(1+exp(z)) in the logit model and it is the cumulative distribution function for standard logistic random variable. The logistic model and linear model are depicted in figure 3.1.

Figure 3.1 Linear and Logit model visualization took from lectures



Source: https://www.researchgate.net/publication/323949660_Comprehensive_Review_On_Supervised_Machine_Learning_Algorithms/figures?lo=1

The logit model derivation. We assume that there is an unobserved (latent) variable y*:

$$y *= \ \beta 0 + \beta 1X2 + \cdots + \beta kXk + e \ and \ y = 1[y *> 0]$$

where βj from LPM and 1[.] is the indicator function that has outcome "1", if y*>0 and "0" if y*≤0. Assume that u has standard logistic distribution and is independent of Xj. Point that 1-G(-z) = G(z)  because e is symmetrically distributed about zero. Thus, the response probability for y:

$$\begin{aligned} P(y=1|X) &= P(y*>0|X) = P(e>\beta 0+\beta 1X2+\cdots+\beta kXk+u|X) \\ &= 1-G(-(\beta 0+\beta 1X2+\cdots+\beta kXk+u)) \\ &= G(\beta 0+\beta 1X2+\cdots+\beta kXk+u) \end{aligned}$$

which is the same function we considered in the beginning.

For partial effects of the logit model, we differentiate the cumulative distribution function:

$$\partial P(X)/\partial Xj \ = g(\beta 0+\beta 1X2+\cdots+\beta kXk+u), where\ g(z)\equiv \partial G(z)/\partial z$$

and after that, we have got a probability density function.

In order to estimate the logit model, we use maximum likelihood estimation (MLE). Let's assume a random sample of size n. The density of Yi given Xi is:

$$\begin{aligned} f(y|Xi;\beta j) = (G(\beta 0+\beta 1X2+\cdots+\beta kXk+u))^{y}(1-G(\beta 0+\beta 1X2+\cdots \\ +\beta kXk+u))^{1-y} \end{aligned}$$

where $y=0\ or\ 1$. When y=1, $f(y|Xi;\beta j)=G(\beta 0+\beta 1X2+\cdots+\beta kXk+u)$ and when y=0, $f(y|Xi;\beta j)=1-G(\beta 0+\beta 1X2+\cdots+\beta kXk+u)$. Now let's take the logarithm and the log-likelihood function for observation i is:

$$li(\beta j)=yi*log(G(\beta 0+\beta 1X2+\cdots+\beta kXk+u)\ +$$

$$+(1-yi)*log(1-G(\beta 0+\beta 1X2+\cdots+\beta kXk+u))$$

The log-likelihood function for the sample n is:

$$L(\beta j)=\sum_{i=1}^{n}\quad li(\beta j)$$

The maximum likelihood estimation of $\beta$is $\hat{\beta} = max(L(\beta j))$. In the logit model, $\hat{\beta}$ is the logit estimator.

In order to find out by how much probability increases or decreases with variable change, partial effects (or margin effects) are used.

# CHAPTER 4. DATA

## 4.1 Descriptive statistics tables

The transaction information about first orders of one Ukrainian startup was taken for the research. First transactions were chosen because they are most critical for business, because for these transactions payment form is filled by users manually. Then payments are performed from a token that was generated after filling the form.

We are dealing with a dataset with 117902 transactions from September 2020 to September 2021. The data was anonymized to prevent the outflow of information about customers and the company. At table 4.1 continuous variables descriptive statistics mean, standard deviation, minimum and maximum are shown.

Table 4.1 Descriptive statistics of variables

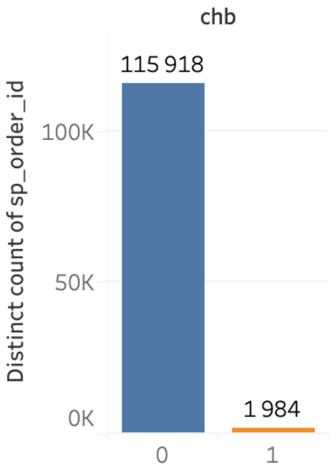
| Variable | Mean | Standard Deviation | Min | Max |
|---|---|---|---|---|
| amount_usd | 1799.32 | 1000.91 | 68 | 5231 |
| real_time_for_form | 88.88 | 72.08 | 1 | 449 |
| true_order_number | 1.13 | 0.46 | 1 | 35 |
| real_cascade_number | 1.10 | 0.34 | 1 | 3 |

## 4.2 Graphics

In this section graphs that represent data in the context of  previously formulated hypotheses are shown and described. In each figure from 4.1 to 4.7 the orange color in the bottom represents the amount of transactions with chargebacks which are considered to be fraud.

At Figure 4.1 we can see the number of non-chargeback transactions on the left, the underwrited "0" and number of chargeback transactions. The chargeback rate is almost 1.6%, which is quite high.

Figure 4.1 Chargeback and non-chargeback transactions



At Figure 4.2 the transaction distribution on price USD (in cents) is depicted. As you can see, there are almost no chargebacks for small amounts. The highest amount of chargebacks are near the 1000 cents mark.

At Figure 4.3 Transaction distribution through the time of users filling the form is shown. The interesting thing is that there are no chargebacks on transactions in which the form was filled in in less than 15 seconds. This may be due to the fact that fraudsters can set the freeze on filling the form for 15 seconds, so as not to be too conspicuous, or data from the cards are entered manually, since there is no information about the card data in the cache of the browser.

Figure 4.2 Transaction distribution on price USD (in cents)

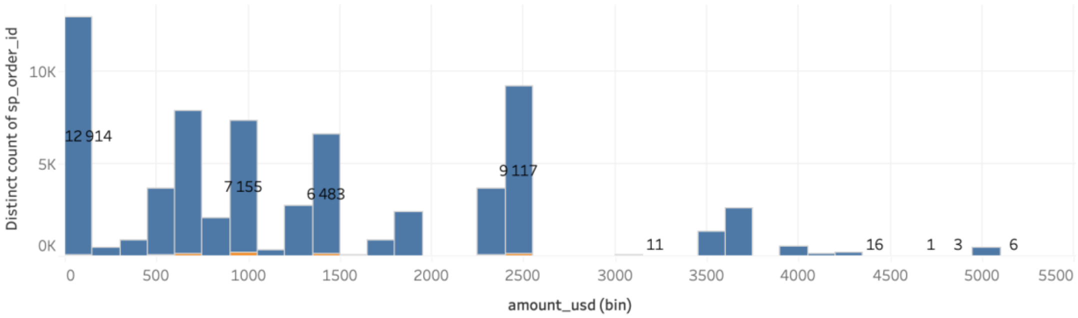
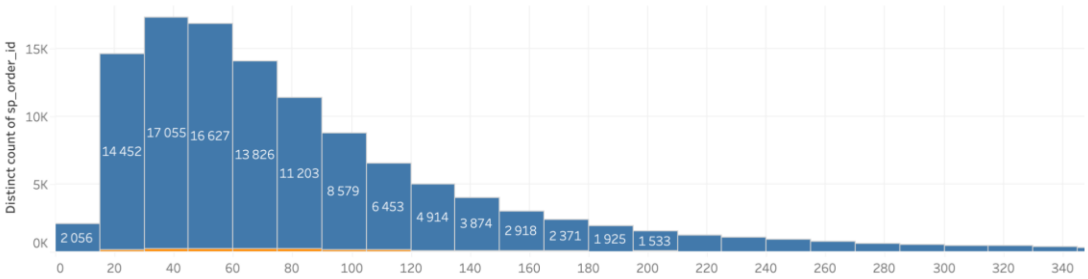


Figure 4.3 Distribution of transactions through the time of users filling the payment form



At Figure 4.4 we can see how chargebacks are distributed among 3D secured transactions and 2D transactions. All transactions with chargebacks were processed without 3D secure protocol.

At Figure 4.5 distribution of transactions by country is depicted. The countries with the most transactions are the USA, Italy, Canada, France, Australia, Russia, the UK, Turkey and Spain. The USA and Canada have the largest number of chargebacks.

At Figure 4.6 we can see how chargebacks are distributed among cardbrands. VISA is the most popular cardbrand among startup customers. The ratio of the number of users of cardbrands in a startup reflects the market share of cardbrands.

Figure 4.4 Chargebacks distributed among transactions with and without 3D secure

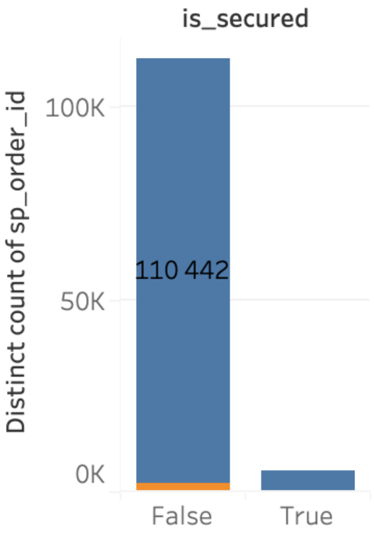


Figure 4.5 Distribution of transactions by country

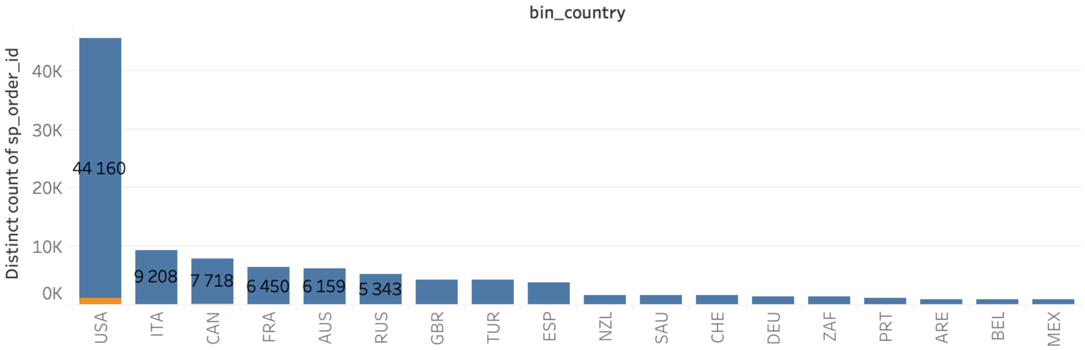


Figure 4.7 shows the distribution of transactions by card type. As you can see from the chart, the most popular type of card for payments is a debit card. The second

most common type is credit cards, but they have more chargebacks than debit cards. Papeid and other cards do not have transactions with chargebacks.

Figure 4.6 Distribution of transactions by cardbrand

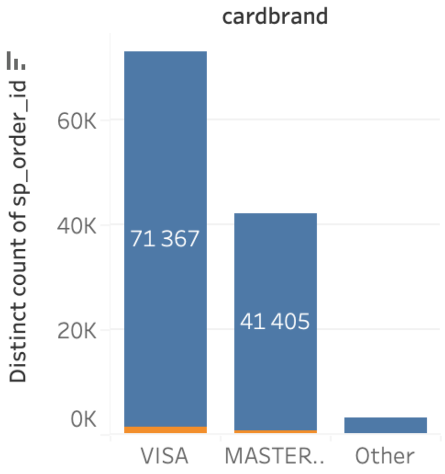
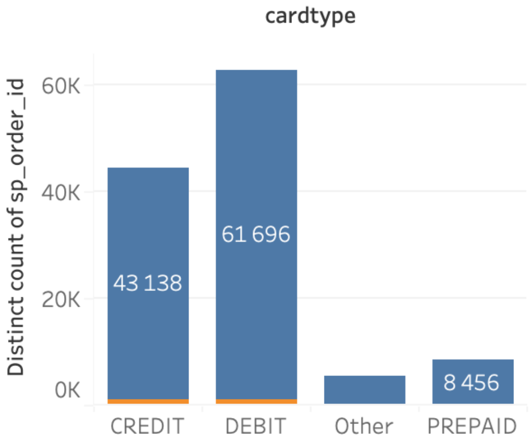


Figure 4.7 Distribution of transactions by cardtype



# CHAPTER 5. RESULTS

The following Logit regression model had been run:

$$CHB = ln(A) + ln(ON) + CN + ln(TF) + D + CM + CBV + CBO +$$

$$+CTD + CTP + CTO + NW$$

Variables were descripted in paragraph 3 methodology. The results of the model are illustrated at table 5.1. Overall the model has LLR p-value = 2.5091e-140. Pseudo R-squared is 0.034 but it is not appropriate to estimate how many observations the Logit model explains with pseudo R-squared. Df Model is 12. Number of iterations is 10 and the number of observations is 117902. Coefficients and their standard errors are shown at table 5.1. Interpretation of betas in the Logit model is not straightforward. Thus we can only rely on the sign of beta and interpret the results.

The interpretation of results are the following:

1) Transactions with higher amounts are more likely to be made by frauders, ceteris paribus. Hypothesis is confirmed;
2) With each additional number of first transaction the probability of its fraudulent increases, ceteris paribus. Hypothesis is confirmed;
3) Cascade number coefficient is not significant, so it can not be interpreted. Hypothesis is rejected;
4) Increase in time the user has spent filling the payment form increases the probability of this transaction being fraudulent, ceteris paribus. Hypothesis is rejected;
5) Transactions which have passed 3D secure are less likely to be fraud, ceteris paribus. Hypothesis is confirmed;

Table 5.1. Coefficient estimation results

| Variable | Coef | Std.Err. | z | P>\|z\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| const | -6.6104 | 0.2921 | -22.6284 | 0.0000 | -7.1830 | -6.0378 |
| amount_usd | 0.2326 | 0.0291 | 7.9883 | 0.0000 | 0.1756 | 0.2897 |
| order_number | 0.2258 | 0.0933 | 2.4208 | 0.0155 | 0.0430 | 0.4086 |
| cascade_number | -0.1218 | 0.0872 | -1.3972 | 0.1623 | -0.2927 | 0.0491 |
| time_for_form | 0.1647 | 0.0312 | 5.2880 | 0.0000 | 0.1037 | 0.2258 |
| is_secured_True | -2.2354 | 0.3056 | -7.3144 | 0.0000 | -2.8343 | -1.6364 |
| country_mismatch_ No mismatch | 0.4727 | 0.1179 | 4.0083 | 0.0001 | 0.2415 | 0.7038 |
| cardbrand_Other | 0.0960 | 0.1354 | 0.7091 | 0.4782 | -0.1693 | 0.3613 |
| cardbrand_VISA | 0.3380 | 0.0513 | 6.5878 | 0.0000 | 0.2374 | 0.4385 |
| cardtype_DEBIT | -0.5504 | 0.0478 | -11.5048 | 0.0000 | -0.6441 | -0.4566 |
| cardtype_Other | -1.3078 | 0.1784 | -7.3304 | 0.0000 | -1.6575 | -0.9581 |
| cardtype_PREPAID | -2.1002 | 0.2078 | -10.1064 | 0.0000 | -2.5075 | -1.6929 |
| written_Real_name | 0.1763 | 0.2361 | 0.7465 | 0.4554 | -0.2866 | 0.6391 |

6) Transactions where user IP country and card issuer county are the same, have higher probability to be fraudulent, ceteris paribus. Hypothesis is rejected;

7) Users with VISA cards are more likely to be frauders than users with MasterCard, ceteris paribus. Hypothesis is confirmed;

8) Other cardbrands coefficient is not significant, so the difference between Mastercard and Other cardbrands (not Visa or Mastercard) in fraud rate is not statistically significant. Hypothesis is rejected;

9) Debit cards are less likely to have fraudulent transactions than credit cards, ceteris paribus. Hypothesis is confirmed;

10) Prepaid cards are less likely to have fraudulent transactions than credit cards, ceteris paribus. Hypothesis is confirmed;

11) Other types of cards (not Credit, Debit or Prepaid) are less likely to have fraudulent transactions than credit cards, ceteris paribus. Hypothesis is confirmed;

12) User card name coefficient is not significant, so it can not be interpreted. Hypothesis is rejected;

In order to know by how much the probabilities are increasing or decreasing, partial effects (the marginal effects) are computed. The results are depicted at table 5.2.

Marginal effect interpretations are the following:

1) One cent increase in transaction amount increases the probability transaction being fraudulent by 2,8%, ceteris paribus;
2) Every subsequent first transaction is 0,03% likely being fraud, ceteris paribus;
3) Cascade_number coefficient is insignificant;
4) Each additional second of filling the payment form increases the probability transaction being fraudulent by 1,7%, ceteris paribus;
5) Transactions with 3D secure has 3% less probability being fraudulent, ceteris paribus;
6) Transactions without county mismatch are 0,7% more likely to be fraudulent than with country mismatch, ceteris paribus;
7) Other cardbrands coefficient is not significant;

8) Visa card transactions have 0,5% higher probability to be fraudulent than MasterCard, ceteris paribus;
9) Debit cardtype transactions are 0,91% less likely to be fraudulent than credit cardtype, ceteris paribus;
10) Prepaid cardtype transactions have 0,3% lower probability to be fraudulent than credit cardtype, ceteris paribus;
11) Other cardtype transactions (not credit, debit or prepaid) are 0,2 % less likely to be fraudulent than credit cardtype, ceteris paribus;
12) User card name coefficient is not significant.

Table 5.2. Logit model one marginal effects

| **Variable** | **Mothod** | **Coef.** | **Std. Err.** |
|---|---|---|---|
| amount_usd | dy/d(lnx) | 0.0284 | 0.004 |
| order_number | dy/d(lnx) | 0.0003 | 0.000 |
| cascade_number | dy/dx | -0.0020 | 0.001 |
| time_for_form | dy/d(lnx) | 0.0117 | 0.002 |
| is_secured_True | dy/dx | -0.0368 | 0.005 |
| country_mismatch_ No mismatch | dy/dx | 0.0078 | 0.002 |
| cardbrand_Other | dy/dx | 0.0016 | 0.002 |
| cardbrand_VISA | dy/dx | 0.0056 | 0.001 |
| cardtype_DEBIT | dy/dx | -0.0091 | 0.001 |
| cardtype_Other | dy/dx | -0.0215 | 0.003 |
| cardtype_PREPAID | dy/dx | -0.0346 | 0.003 |
| written_Real_name | dy/dx | 0.0029 | 0.004 |

# CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

In the research patterns of fraudulent transactions performances in Ukrainian startup company were explored. The Ukrainian startup industry and online payments industry, its recent trends were described, the related studies were overviewed. 12 hypotheses of fraud payment patterns and performance were formulated, methodology and the Logit model were described. Nine out of twelve coefficients are statistically significant.  The data was collected, visualized and described.

Ukrainian startup industry is growing rapidly and the startup ecosystem in the country and especially in the capital Kyiv is on a decent level and  has been improving in recent years. That opens new horizonts to entrepreneurs and gives opportunities to succeed.

Many startups focus on selling their products or services online. To do this, they need to establish a system for accepting online payments. One of the problems that startups face is the high rate of fraud and cybercrime. Based on the work done, recommendations can be formulated for startups to identify and suppress fraudulent transactions.

Based on the results of the research, the conclusions and recommendations for businesses which perform online are the following:

1) Make high amount transactions more secured.

High price products are usually the target for fraud. It is a good idea use 3D secure or other extra types of verification for such kind of transactions. This should not affect conversion, because users usually are more attentive to high price purchase, but will significantly decrease amount of loss caused by fraud.

2) Pay attention to a large number of attempts to make the first payment or a large number of first payments on one account.

A large number of unsuccessful transactions or several successful first orders from one account may be a sign of fraud. Most likely, the fraudster wants to check for which cards he has the correct credentials or to empty as many cards as possible for a short period of time.

3) You can fearlessly send payments to the next steps of the cascade.

If the payment did not go through the first step of the cascade and switch to another, this does not mean that the payment is more likely to be fraudulent. This research showed that this coefficient is insignificant. You can successfully increase your conversion without a fear of accepting fraudulent payments.

4) Quick filling of the payment form does not always signal that the user is a fraudster.

Fraudsters are constantly improving their approaches in order to not be conspicuous. Most likely, the fraudsters could set the freeze for program that fills the payment form, so they did not fill it quickly. Regular users can now also fill out payment forms fast because of the browser hash. Or large time of filling the form could be one of the indicators of friendly fraud.

5) Use 3D secure to protect from fraudsters.

3D secure has shown its effectiveness in preventing fraudulent transactions. We recommend to find a balance between 3D secure and high conversion.

6) Pay more attention to payments with VISA cards.

The research showed that fraudulent transaction rate is mote via VISA cards than MasterCard or other cards. Since VISA cards are the most popular, fraudsters possible tend to deal on them the most.

7) Pay more attention to transactions made by credit cards.

Credit cards showed a higher level of fraud than other types of cards.

In the future this research can be developed into a machine learning antifraud system that will help companies detect and reject fraudulent transactions. Also, new hypothesis can be formulated from a broader set of data. For example, data on the machine fingerprint of the device from which the transaction was requested, can greatly help in the detection of industrial fraud.

# REFERENCES

Chalfant James, and Alston Julian. 1988. *Accounting for Changes in Tastes. Journal of Political Economy*

Deaton Angus, and J. Muellbauer. 1980. *An Almost Ideal Demand System. American Economic Review*

Fast Market Research. Ukraine Agribusiness Report Q3 2015. Published on May 7, 2015.
http://www.fastmr.com/prod/991633_ukraine_agribusiness_report.aspx?afid=101

Golan Amos, Perloff Jeffrey, and Shen Edward. 2001. *Estimating a Demand System with Nonnegativity Constraints: Mexican Meat Demand. The Review of Economics and Statistics*

State Statistics Service of Ukraine. Statistical information.

http://www.ukrstat.gov.ua.

Sunil Erevelles, Nobuyuki Fukawa, Linda Swayne. 2015. *Big Data consumer analytics and the transformation of marketing. Journal of Business Research*

Jeffrey M. Wooldridge. 2019. *Introductory Econometrics: A Modern Approach. Cengage Learning*

Babiachok R., Kulchytsky I. 2018 *Analytical material. Main Trends Of Startup Development In Ukraine Problems, Obstacles And Opportunities.*

A. Brabazon, J. Cahill, P. Keenan, D. Walsh, *Identifying online credit card fraud using Artificial Immune Systems*

Ohad Samet, June 2013*, Introduction to Online Payments Risk Management*
https://www.civic-synergy.org.ua/wp-content/uploads/2018/04/Osnovni-tendentsiyi-rozvytku-startapiv-v-Ukrayini-1-1.pdf.

Startup Ranking. Statistical information.

https://www.startupranking.com/countries.

Startup Blink. Startup Ecosystem Rankings 2020. https://drive.google.com/file/d/1ydvc0_jA3g8fC07vzfN8GWqEkP6yQ_Ak/view?mc_cid=51b334201c&mc_eid=561226bfc3.

Official website of Ukraine, administered by the Ministry of Foreign Affairs of Ukraine and Ukrainian Institute. UKRAINIAN STARTUPS. https://ukraine.ua/invest/startup/.

Cornell University, INSEAD, the World Intellectual Property Organization. *Global Innovation Index 2020, 13th edition.* https://www.wipo.int/edocs/pubdocs/en/wipo_pub_gii_2020/ua.pdf.

Logistic Regression figure. https://www.saedsayad.com/logistic_regression.htm.

Digital Payments Market by Component (Solutions (Payment Processing, Payment Gateway, Payment Wallet, POS Solution, Payment Security and Fraud Management) and Services), Deployment Type, Organization Size, Vertical, and Region - Global Forecast to 2025 https://www.marketsandmarkets.com/PressReleases/digital-payment.asp