PAYMENT FORM CONFIGURATION AND APPROVAL RATE OF TRANSACTIONS IN ONLINE PAYMENTS

by

Maksym Chvartkovskyi

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

MA in Business and Financial Economics

Kyiv School of Economics

2024

Thesis Supervisor: ______________________________Professor Olena Besedina_

Approved by

Head of the KSE Defense Committee, Professor

Date ______________________________________

ACKNOWLEDGMENTS

I would like to thank my thesis advisor, Professor Olena Besedina, for the guidance and support throughout this journey. Her insights and encouragement were crucial for the end result.

I am also grateful to all the professors at KSE for making last year and a half a wonderful journey of learning economics. The academic satisfaction of the learning process is the one I could not imagine.

I am thankful to the KSE Foundation and Ampersand Foundation for their financial help and mentorship. And special thanks go to my friends and family for their constant support.

Lastly, I want to take a moment to appreciate my classmates and TAs, whose friendship and experience share made this process unforgettable. The real prize is the friends we've made along the way.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

FinTech Financial Technology

PCI DSS Payment Card Industry Data Security Standard

PSP Payment Service Provider

AR Approval Rate

SR Submit Rate

OS Operational System

USD United States Dolar

CAGR Compound Annual Growth Rate

CPA Cost Per Acquisition

MLE Maximum Likelihood Estimation

OLS Ordinary Least Squares

LPM Linear Probability Model

VIF Variance Inflation Factor

CHAPTER 1. INTRODUCTION

In the disrupting phase of financial technology development, the payments sector has emerged as a critical component of the digital economy. As businesses increasingly shift their operations online, the efficiency and effectiveness of payment systems have become very important. One crucial metric that significantly impacts both merchants and consumers is the payment approval rate – the percentage of attempted transactions that are successfully processed.

Despite its importance, the relationship between user behavior patterns on payment forms and subsequent approval rates remains an understudied area in public research. This thesis aims to fill this knowledge gap by examining how various user interactions with payment interfaces correlate with the likelihood of transaction approval. Payment approval rates are a key performance indicator for online businesses across all sectors. A high approval rate indicates a smooth transaction process, leading to increased customer satisfaction and, ultimately, higher conversion rates. Conversely, low approval rates can result in lost sales, frustrated customers, and potential damage to a company's reputation.

Every declined transaction represents a lost opportunity – a customer who was willing to pay for a product or service but was unable to complete the purchase. This scenario is also frustrating for businesses, as it means that despite successfully marketing their product and convincing a customer to pay for it, they fail at the final stage of the sales funnel. In an era where customer retention is critical for stable growth, ensuring a frictionless payment process is crucial for maintaining a loyal customer base.

Considerable research has been conducted on different aspects of e-commerce and online payments, but still there is a lack of publicly available studies focusing specifically on the relationship between configuration of payment forms and approval rates. Several factors may contribute to this knowledge gap. Payment processors and large e-commerce platforms likely possess big amounts of data on user behavior and approval rates. However, this information is often considered proprietary and is not shared publicly, as it can provide a significant competitive advantage. The sensitive nature of payment data and strong regulations surrounding financial information make it challenging to conduct and publish comprehensive studies in this area. The fast-paced development of payment technologies means that user interfaces and backend systems constantly evolve, potentially making older studies less relevant. The number of factors affecting payment approval rates, including card issuer policies and fraud detection algorithms, adds a lot of complexity to any analysis.

The primary objectives of this research are following. First of all, to analyze the prospects of online payments industry and review the related to the topic literature. Secondly, to design a methodology and collect the data to identify the relationship between payment form configuration and user's behavior to approval rate in online transaction . Thirdly, to present the results of the research in a simple and easy to understand manner. Lastly, to propose strategies for optimizing payment forms and processes to improve approval rates based on the findings.

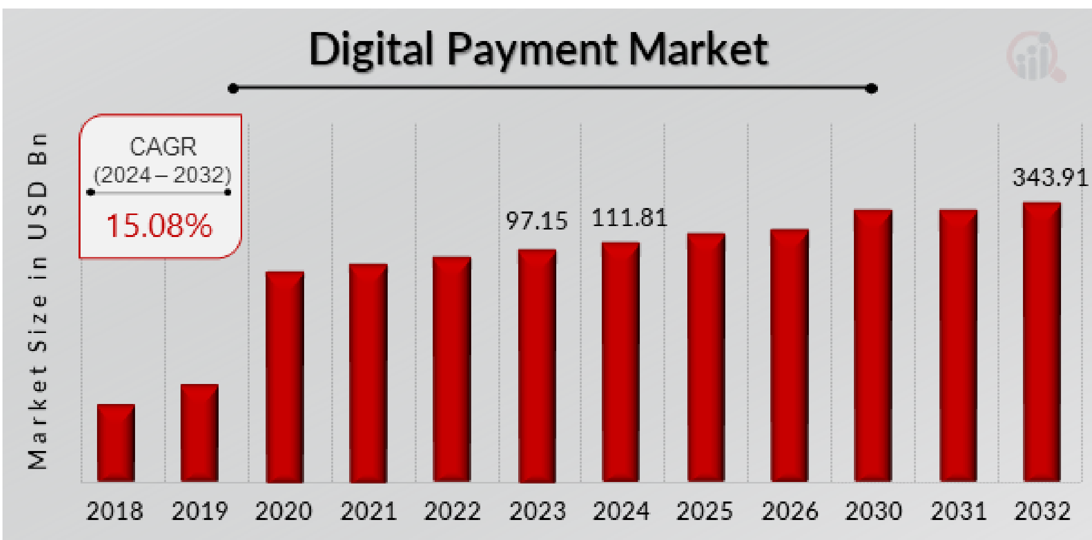
Chapter 2 of the thesis gives an overview of how fintech experts and reputable journals estimate the potential growth of online payments by cards. Also, this chapter reviews the literature, related to the topic, even though, there are not much directly related publicly available sources. Chapter 3 explains the methodology behind the quantitative research and provides the proposed model and variables to examine the relationships between configuration of the payment form and user's behavior to the approval of a payment. Chapter 4 presents the descriptive statistics and insights about data for the study. It also covers the basic visual representation of a relationship between examined variables. Chapter 5 compares results of two estimations of the model and provides model validation using different methods and interpretation of the statistics software output. Chapter 6 provides the suggestions that can be incorporated by businesses to utilize the results of this research.

CHAPTER 2. INDUSTRY OVERVIEW AND RELATED STUDIES

2.1 Industry overview

The fintech payment industry has experienced explosive growth over the past decade, revolutionizing the way individuals and businesses conduct financial transactions. This rapid expansion has been driven by technological advancements, changing consumer preferences, and the increasing digitization of commerce. The global digital payments market has seen consistent year-over-year growth since the early 2010s. This growth trajectory is expected to continue, with Compound Annual Growth Rate (CAGR) projections for the online payments industry ranging from 15% to 21% over the next 5-8 years. This robust growth forecast underscores the increasing importance and adoption of digital payment solutions worldwide.

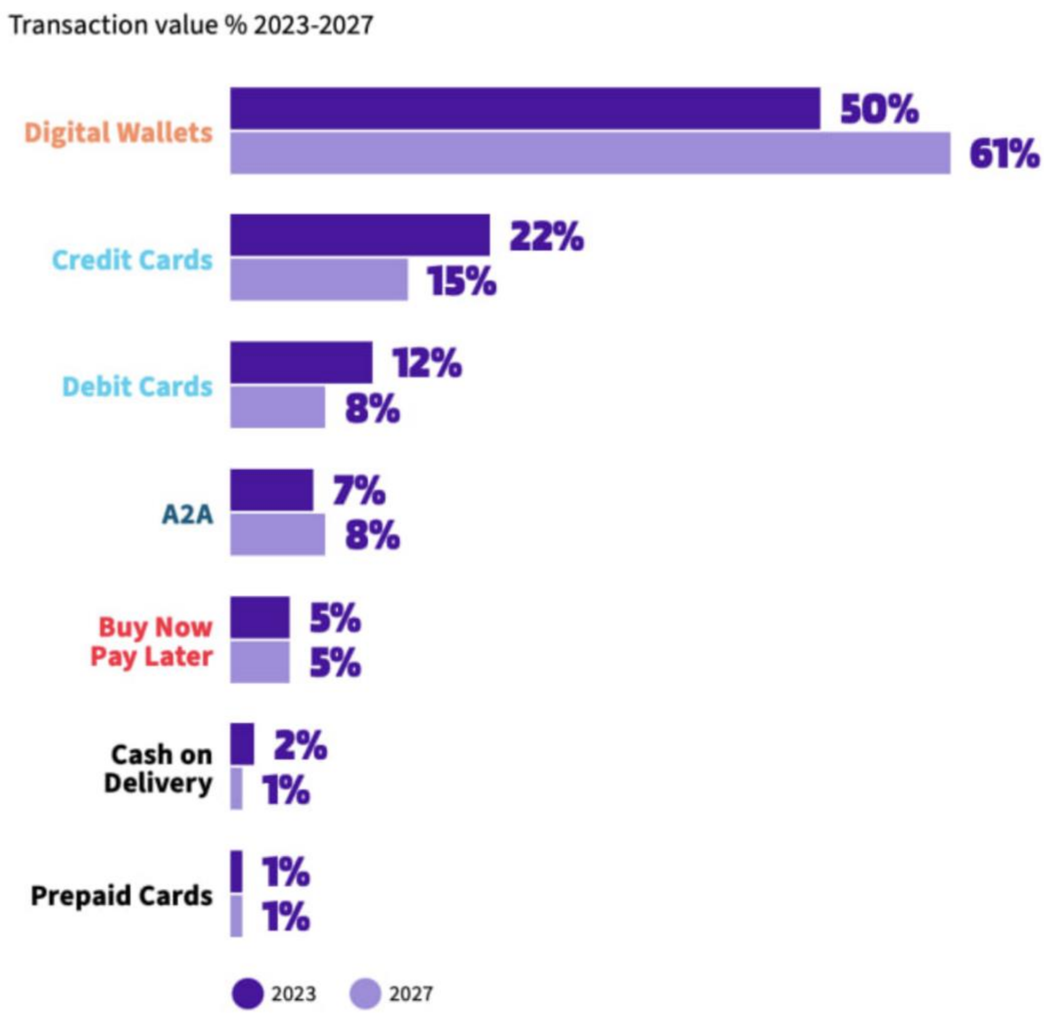
Figure 2.1. Digital Payment Market Size, 2024-2032 (USD Billion)



Source: Aari Dhaptel, 2024

While the overall digital payments sector is expanding, it's important to note the evolving landscape within the industry. Card payments, which have traditionally dominated the market and are the focus of this research, are facing increasing competition from alternative payment methods. In particular, digital wallets are gaining market share at the expense of card payments. However, despite this shift, industry analysts predict that debit and credit card payments will remain relevant and continue to play a significant role in the payments ecosystem for the foreseeable future. In major markets such as the European Union and North America, card payments still constitute a substantial portion of the payment landscape as presented on the Figure 2.2.

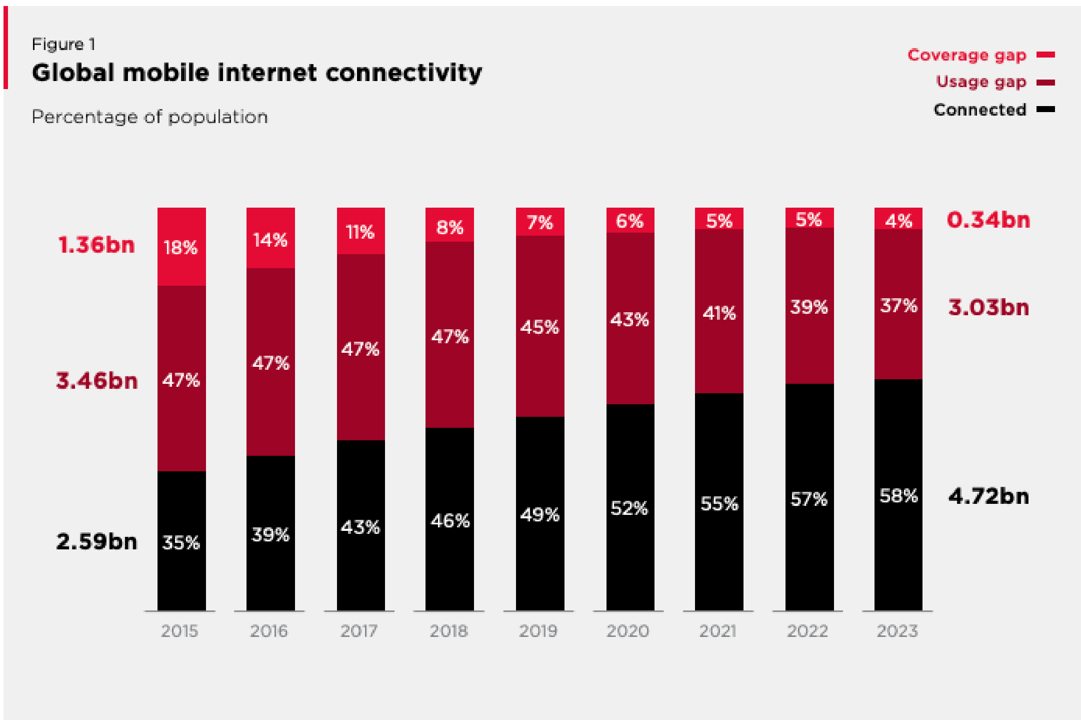
Figure 2.2. Global e-commerce payment methods share, %



Source: Worldpay, "*Global Payments Report 2024*"

There are a number of factors that contribute to the payments sector growth. Smartphone and mobile internet penetration. Since 2015, number of users connected to mobile internet almost doubled: from 2.6 billion users in 2015 up to 4.72 billion users in 2023 as you can see in Figure 2.3. This adoption of mobile technologies provided a platform for users fro mobile payment solution. It did make online financial services accessible and easier to use.

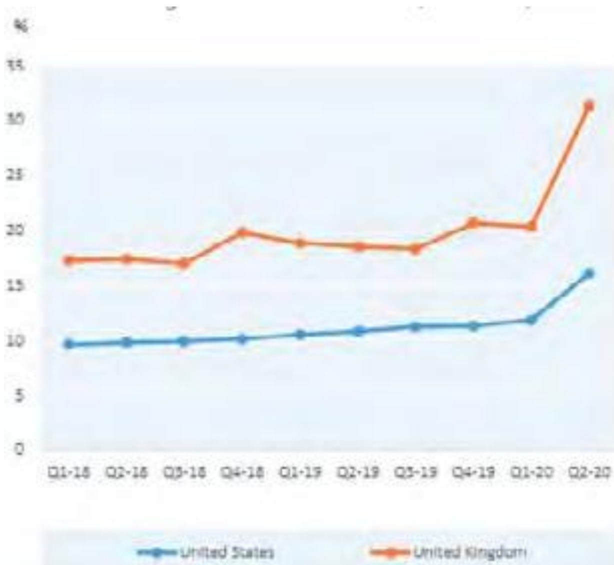
Figure 2.3. Global Mobile internet connectivity



Source: GSMA, "*The mobile economy 2024*"

The rapid growth of online shopping has necessitated robust digital payment systems, driving innovation in the sector. At the same time the global health crisis during COVID-19 pandemic has accelerated the shift towards digital payments as consumers and businesses sought contactless transaction methods. Rapid increase in share of e-commerce retail sales in United Kingdom and United States, that are amongst the world's largest digitalized markets is presented on Figure 2.4

Figure 2.4 Share of e-commerce in total retail sales, UK and US (2018-2020)



Source: UN, "*COVID-19 and E-commerce*"

As the industry has matured, leading payment service providers have recognized the importance of offering comprehensive solutions that go beyond just transaction processing. One significant trend in this matter is providing free value-adding services such as hosted payment pages or customizable payment forms.

These services offer several advantages. By handling payment data directly, providers can ensure compliance with the latest security standards, including PCI DSS (Payment Card Industry Data Security Standard). This is particularly beneficial for merchants lacking the resources or expertise to maintain compliance by themselves. Specialized payment providers can leverage their expertise to optimize payment interfaces, potentially increasing conversion rates and reducing cart abandonment. When using a hosted payment page, merchants do not directly handle sensitive payment information, thus reducing their liability in case of data breaches. These solutions are typically designed to work across various devices and platforms, saving merchants the effort of developing and maintaining multiple payment interfaces. Advanced fraud detection mechanisms can be built into these payment pages, leveraging the provider's extensive data and expertise to minimize fraudulent transactions.

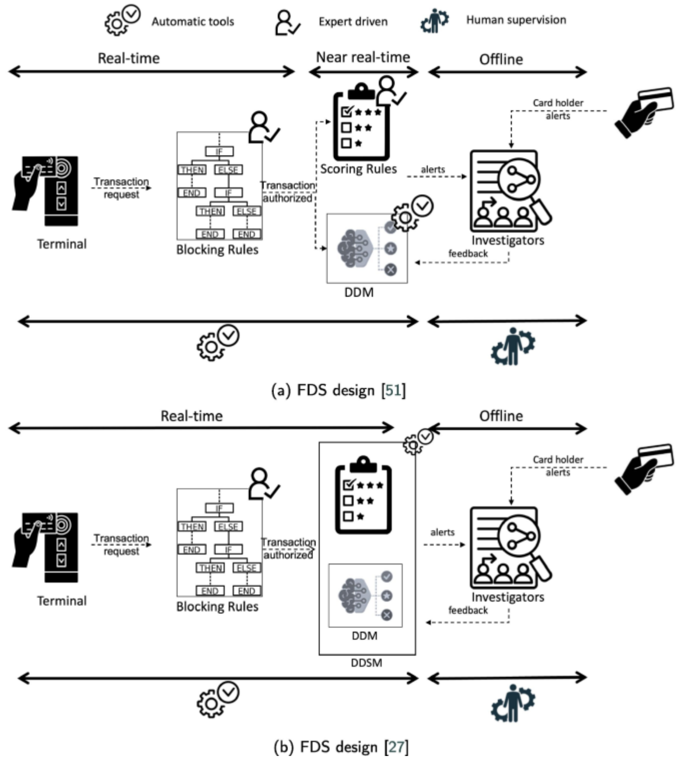
Despite its rapid growth, the fintech payments industry faces several challenges. Navigating complex and evolving financial regulations across different jurisdictions remains a significant challenge. As payment systems become more digital, they also become targets for cybercriminals, necessitating ongoing investment in security measures. Looking ahead, the industry is likely to see continued growth and innovation. Emerging technologies such as blockchain, artificial intelligence, and the Internet of Things are expected to play increasingly important roles in shaping the future of payments.

2.2 Related studies

Related studies mainly cover the topic of payment fraud prevention. Payment fraud prevention is the reason for the number of card payment declines that are aimed to be decreased with recommendations from this study.

Payment fraud detection works as a combination of real-time scoring models and risk-rules activation and hands-on investigation by experts in this field (Dal Pozzolo et al., 2018). This enables fast yet effective protection of payers from the payment fraud.

Figure 2.5 Fraud detection system design



Source: Cherif et al., 2023

The challenge for the merchants hides in details: they need to provide all the necessary data and try to push customers into the most effective flow to maximize the probability that the payment will successfully pass the issuer's antifraud system.

Yet, it is crucial not to let fraudulent transactions pass the payment flow to the authorization step. Because it may negatively influence the issuer's risk-scoring model.

CHAPTER 3. METHODOLOGY

The difference between approval rate of different cohorts of online transactions can be associated with the differences in configurations of payment form and user behaviour on the payment form during the payment. This is observable characteristics that we can incorporate in our model of interest. On the other side, there are number of unobservable characteristics that can not be measured in this research. This means that some of the volatility in an dependent variable can not be captured. Examples of such unobserved characteristics are marketing spend, cost per click, brand index, historical processed volume of a merchant etc.

For the investigation of the impact of an observable characteristics on a dependent variable, which is a binary variable of approval of an online transaction, research uses the Probit model using maximum likelihood estimation. The estimated models and detailed explanation for every exogenous variable are presented below.

3.2 Model and variables

$Approve = \beta_0 + \beta_1 is\_cardholder\_required + \beta_2 is\_postal\_code\_required + \beta_3 amount + \beta_4 number\_autofilled\_fields + \beta_5 \ln(time\_to\_submit) + \beta_6 number\_errors\_occurred + \beta_7 number\_tooltips\_hovered + D\gamma + \varepsilon$

Where $D\gamma$ is a vector of control dummy variables and following variables are incorporated:

is_cardholder_required – dummy variable that indicates that the cardholder name was required on payment form. Card holder name is a field that may signalize the issuer that the transaction is initiated by a genuine customer.

is_postal_code_required – dummy variable that indicates that the postal code was required on payment form. Billing address postal code is commonly a five-digit number that is dealt with as a knowledge that may be known only to the real owner of a card, because it is not written on the card itself. Even though it is not a kind of password and can be stolen, it is a way for issuer to verify more information in a processed transaction.

number_autofilled_fields – variable that represents the number of fields on payment form that were inserted using the autofill. Autofill occurs when users of a browser has already entered the payment data using this particular browser and accepted the browsers proposition to save card information. This way, next time a user encounters the payment form on any other website, he or she can autofill the payment details. It is a valuable information, collected specifically for this research. It is important because it critically influences the way users interact with the payment form: from manually entering all the data to simple click.

time_to_submit – time that takes for a user from seeing a payment form to submitting the payment. This variable is also specifically tracked for the sake of this research. It can capture the carefullness of the user while entering the payment details, whilst also considering the experience of the users: more experienced users are expected to submit payment form faster.

amount – amount of payment in USD cents. For payment that are not charge in a United Stated dollar, it is converted automatically at the exchange rate at the moment of a transaction.

is_3Ds – control dummy variable included in $D\gamma$ that indicates that the payment received 3Ds verification. 3Ds verification is a technology to provide authentication of a particular user. A common way to perform an authentication is by using push messages in banks mobile app or sending a message with one time password to the payer's mobile phone.

is_country_matched – control dummy variable included in $D\gamma$ that indicates that the country interrogated by IP address matches the country interrogated from card bin. Card bin is a first six or eight digits in a payment card number. It indicates the information about the card, specic to where and how it was issued: issuing country, bank, type of card, cardbrand. Intuition behind this variable is that ordinarily people tend to pay from the same country where the country was issued. So mismatch between IP country (which is considered a real country of users presence) can be associated with different user or user using VPN. Both of this option are not favorable by issuer.

is_exotic_browser – control dummy variable included in Dγ that indicates that the browser of payment is in bottom 5% of popular browsers. Popular browser are found empirically in a dataset, this are the browsers that most of the users are using and that are best tested by the merchants and payment providers. Unpopular browsers can witness poor performance of payment form as a web product and subsequently lower approval rates.

is_exotic_cardbrand – control dummy variable included in Dγ that indicates that cardbrand of a payment card was not "Visa" nor "Mastercard";

number_error_occurred – variable that represents the number of times an errors occurred on payment form;

number_tooltip_hovered – variable that represents the number of times a user have read an advice for some form field;

is_tracking_disabled – control dummy variable included in Dγ that indicates that the users has some blocking plugin on his browser that disables tracking of actions;

is_mobile – control dummy variable included in Dγ that indicates that payer uses the mobile device;

is_desktop – control dummy variable included in Dγ that indicates that payer uses the desktop device

The concise and brief description of variables and their expected signs are presented in the following table

Table 3.1 Variable descriptions and expected signs

Variable	Symbol	Expected sign
Number of autofilled fields	number_autofilled_fields	-
Log of time to submit	time_to_submit	+
Dummy for 3Ds	is_3Ds	+
Dummy for cardholder name required on form	is_cardholder_required	+
Dummy for postal code required on form	is_postal_code_required	+
United States dollars amount	amount_usd	-
Dummy about match of BIN and IP countries	is_countries_matched	+
Dummy for exotic browsers	is_exotic_browser	-
Dummy for exotic cardbrand	is_exotic_cardbrand	-
Number of errors occurred	number_errors_occurred	-
Number of tooltips hovered	number_tooltips_hovered	+
Dummy for disabled tracking	is_tracking_disabled	-
Dummy for mobile platform	is_mobile	-
Dummy for desktop platform	is_desktop	+
Dummy for Windows OS	is_Windows	-
Dummy for iOs OS	is_iOs	+
Dummy for Android OS	is_Android	-

Ideally, we would need to also gather some characteristics about users itself. Such as gender, age, education, as this would provide more holistic view on control variables and isolate this factors from the final equation. Also the marketing data on how the user was acquired would be very valuable. Because if the cost per acquisition of the user is high, we would expect the user to have more chances of approval of a transaction. For those users, whose acquisition was relatively cheap, low approval rate is not a problem, rather just a point of optimization.

3.3 Estimation approach

The study takes into account two types of estimation that can be performed on the model: ordinary least squares estimation utilizing the linear probability model and the maximum likelihood estimation utilizing the probit model. Since the dependent variable is dealed as a binary variable: the customer either succeed in the payment or not, this means that linear probability estimation presents a number of caveats. Firstly, it does not restrict probability to lie between 0 and 1. Even though probability beyond this numbers makes no sense in interpreting. Also, the linear probability model can yield heteroskedasticity and so bias in the estimation results. Furthermore, it assumes a linear relationship between endogenous and exogenous variables, which is probably nor accurate for the case.

Even though linear probability model is much easier to interpret, it is decided to utilize the Probit model using the maximum likelihood estimation. The interpretation of this model may be a challenge because we can not state the one and only effect that the independent variable will imply on the dependent one. However, it is better to be roughly right than precisely wrong.

While developing a model, I needed to consider the risks that I need to overcome to keep the model as precise as possible. In order for the model to not create a biased reflection of the real world, it should be restricted from multicollinearity and heteroskedasticity. In this subsection I will suggest the proposed solution on how to overcome this challenge.

To overcome the multicollinearity, this research covers examination of the data. The independent variables need to be uncorrelated and fully independent from one another. We will come back more to this matter while describing the data that are gathered for the research in Chapter 4. Also, the model needs to be tested on multicollinearity by itself after the esimation. To do this, the Variance Inflation Factor (VIF) will be used to quantify the level of multicollinearity between the independent variables. Variables with VIF values greater than 5 will be carefully examined and potentially adjusted or removed to improve model reliability.

CHAPTER 4. DATA

This section describes the data selection and descriptive statistics of what is used to regress the model presented in a "Methodology" chapter.

For purposes of this research, the data of a few online businesses with a web-to-app funnel was taken. The data consists of first payments, which later may be transformed into subscriptions. Subscription plans make the approval rate metric even more crucial because they ensure that not only the payment amount but also the whole lifetime value of a user on the product is at stake.

Also, in this research, we focus on card payments because digital wallets and alternative payment methods form filling is standardized and can not be customized by merchants or payment providers. Hence, there would be no possibility to utilize the results of the study.

We are looking at dataset that consists of all the necessary data points for about 1.1 million transactions from August 2024. Data is anonymized, so there are no security issues or proprietary data leaks.

Descriptive statistics about all the variables in the dataset are provided in the table below.

Figure 4.1 Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
id	1,123,728	1,047,968,889.000	5,238,912.000	1,038,819,490	1,056,920,587
is_approved	1,123,728	0.658	0.475	0	1
is_3ds	1,123,728	0.163	0.369	0	1
is_cardholder_required	1,123,728	0.450	0.497	0	1
is_postal_code_required	1,123,728	0.354	0.478	0	1
amount_usd	1,123,728	1,468.406	1,754.760	0	147,453
is_country_matched	1,123,728	0.859	0.348	0	1
is_tracking_disabled	1,123,728	0.019	0.137	0	1
time_to_submit	1,123,728	97,355.590	108,651.900	1	23,882,276
number_autofilled_fields	1,123,728	0.848	1.351	0	14
number_error_occurred	1,123,728	0.165	0.767	0	66
number_tooltip_hovered	1,123,728	0.030	0.255	0	93
is_mobile	1,123,728	0.815	0.389	0	1
is_desktop	1,123,728	0.153	0.360	0	1
is_exotic_cardbrand	1,123,728	0.049	0.217	0	1
is_iOS	1,123,728	0.338	0.473	0	1
is_Android	1,123,728	0.489	0.500	0	1
is_Windows	1,123,728	0.061	0.240	0	1
is_MacOs	1,123,728	0.034	0.180	0	1
is_exotic_browser	1,123,728	0.072	0.259	0	1

Overall approval rate in this dataset is 65,8% which is quite normal in online industry. However, it still needs improvement.

Figure 4.2 illustrates the distribution of transactions by USD amount, with each bar’s height representing the number of transactions within a specific 2.5 USD bin. The line graph and color gradient show the approval rate for each bin, indicating that both very small and very large transactions have lower approval rates. This trend may reflect issuers' antifraud measures, as financial institutions often view atypical transaction sizes as higher risk. Small transactions might raise suspicion of fraudulent testing activity, while large transactions could be flagged for exceeding typical spending limits, potentially seen as an indicator of fraud. The decline in approval rates for transactions over 40 USD aligns with these risk-based rules, suggesting a cautious approach from issuers when handling amounts outside an expected transaction range.

Figure 4.2 Distribution by amount

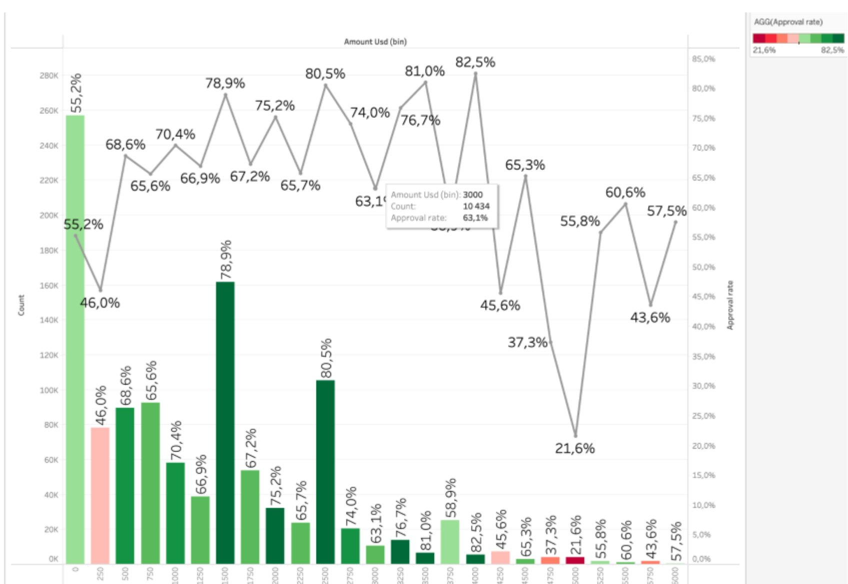


Figure 4.3 displays the distribution of payments and approval rates based on the number of autofilled fields, with bar height reflecting transaction counts. When no autofill is used

(0 fields), the approval rate is high, possibly because users who enter data manually are more cautious. Approval rates drop sharply for transactions with just one autofilled field, likely because it's an uncommon way to fill forms and has fewer observations. As autofill increases to 2 or more fields, approval rates show a steady decline, suggesting issuers may view higher autofill usage as a potential risk. Transactions with 4 or more autofilled fields are rare, making them stand out as possible outliers.

Figure 4.3 Autofill distribution

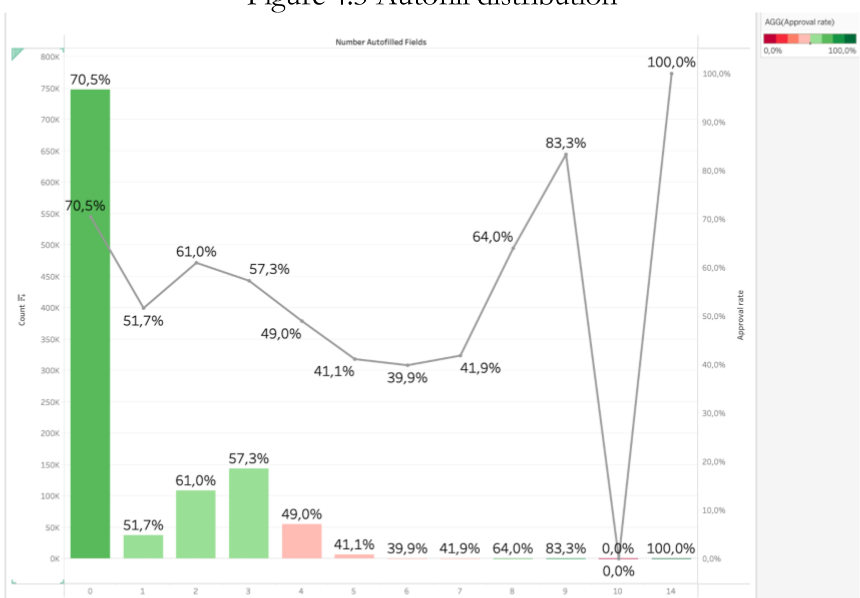


Figure 4.4 shows how transaction approval rates change with the number of errors users encountered while filling the form. Approval rates are highest when no errors occur (66.7%) but drop steadily as errors increase. This trend might indicate that users struggling with initial validation steps, like the LUHN algorithm, are more likely to face further errors. Error counts above 4 are rare and are treated as outliers, with some unusual peaks showing 100% approval likely due to very low observation counts.

Figure 4.4 Distribution by errors encountered

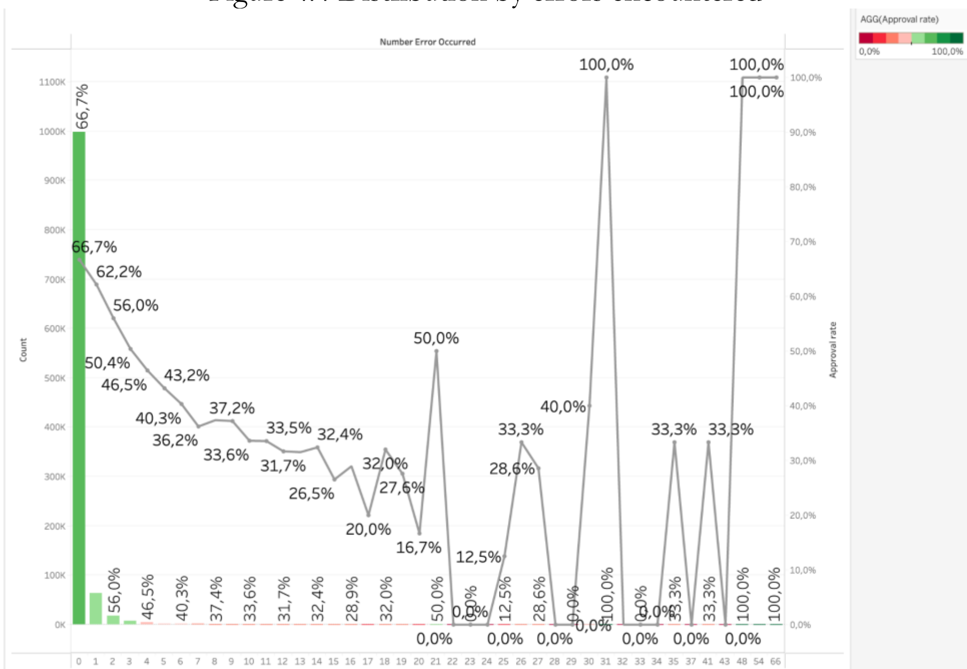
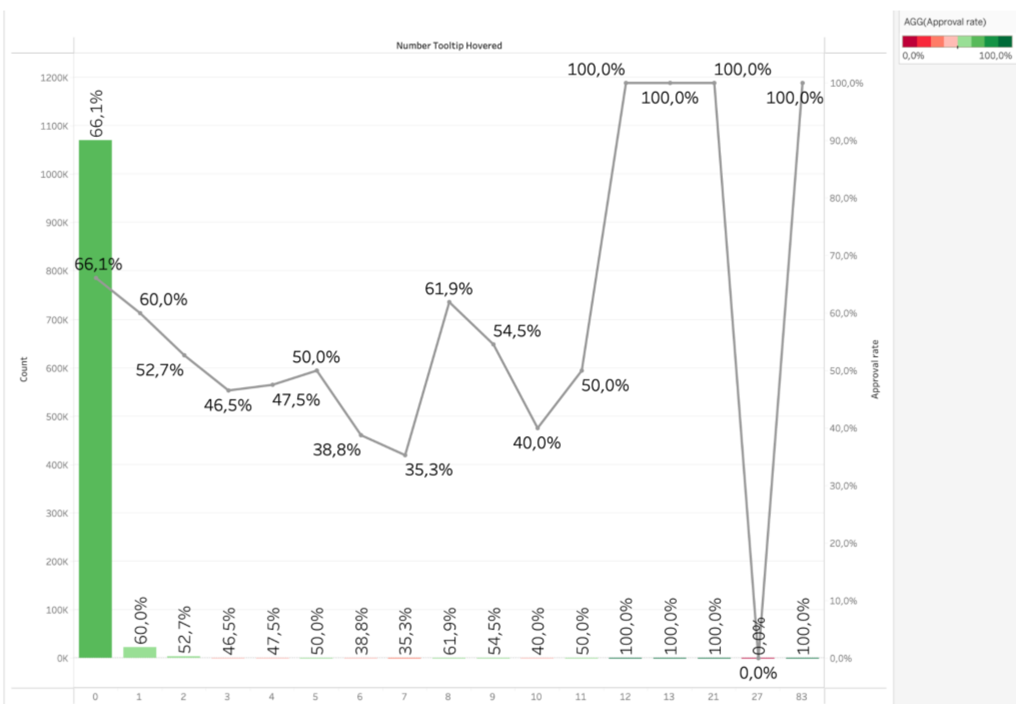


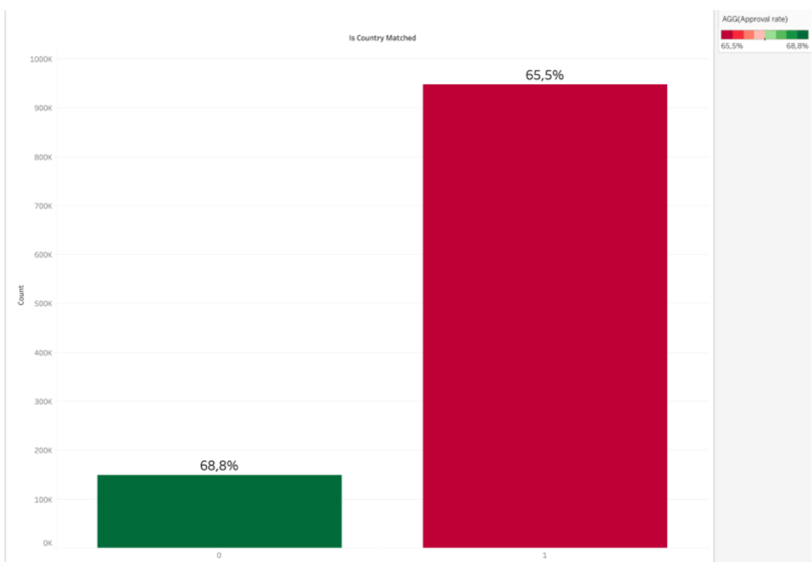
Figure 4.5 Distribution by tooltips hovered



We observe negative correlation between this numbers and approval rate. However number of observations is quite low, so this correlation should be tested.

Figure 4.6 illustrates the relationship between the approval rate and whether the issuing country of a payment card matches the country identified by the IP address. Surprisingly, the approval rate is slightly higher (68.8%) for transactions where the country does not match compared to when it does (65.5%). This unexpected result suggests that other factors might be influencing approvals in the "not matched" group, potentially offsetting typical risk assumptions about mismatched locations.

Figure 4.6 IP country and issuing country match



Figures 4.7 and 4.8 depict distributions of payments and their AR based on characteristics of device and operating system that users utilize.

Figure 4.7 shows the distribution of transactions by operating system and their respective approval rates. Among the five most common OS, Android has the highest usage but a moderate approval rate of 57.3%. iOS, with a higher approval rate of 72.4%, may attract users with greater buying power due to the generally higher cost of iOS devices. Windows and Linux show approval rates of 67.6% and 84.9%, respectively, with Linux potentially appealing to more experienced users. Mac OS, with fewer users, has a moderate approval rate of 75.7%.

Figure 4.7 Distribution by operating system

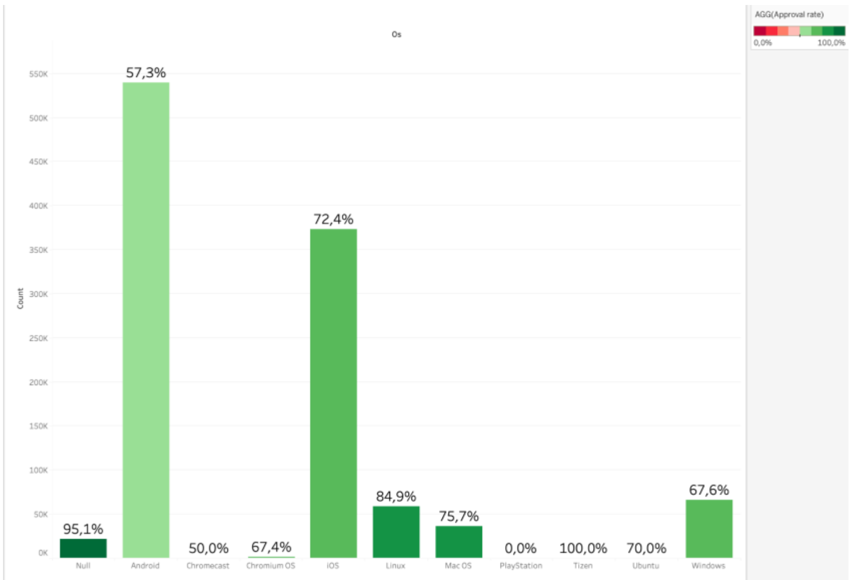
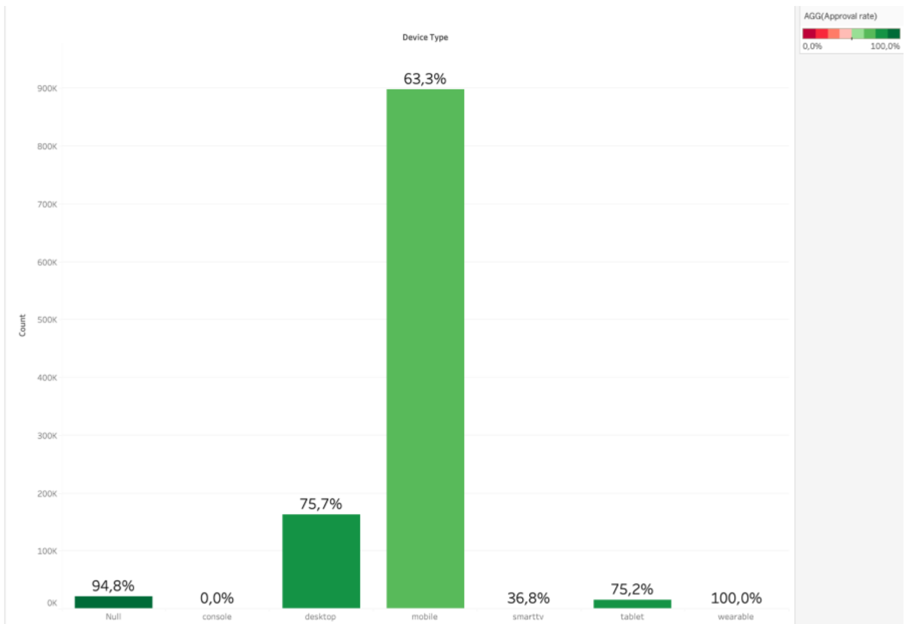


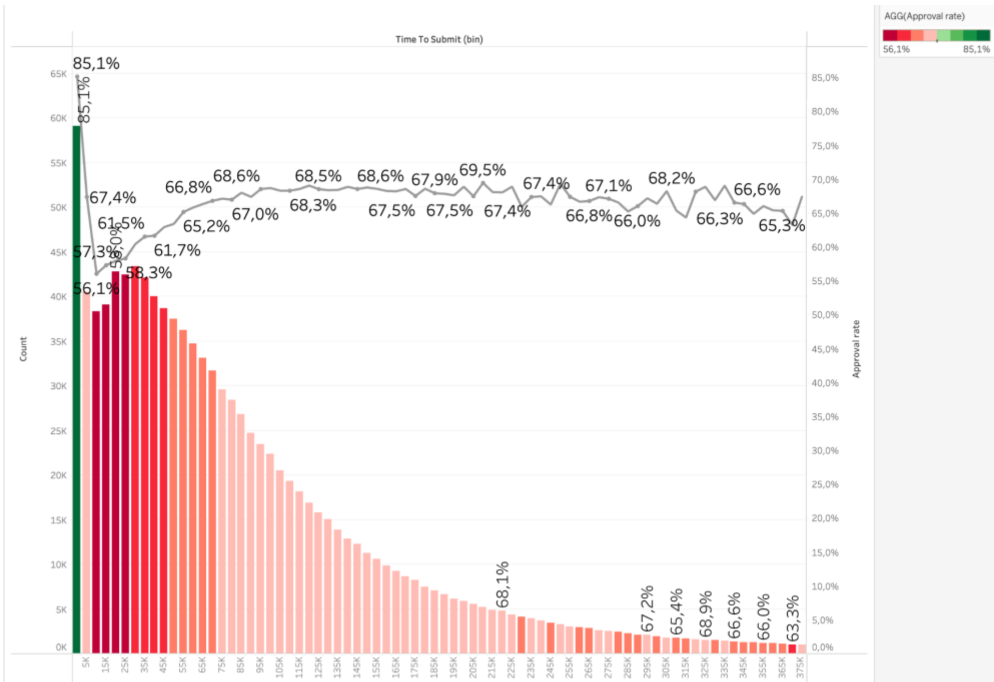
Figure 4.8 shows transaction distribution and approval rates by device type. Mobile devices are the most commonly used, with an approval rate of 63.3%. Desktops have a higher approval rate of 75.7%, suggesting they are seen as more reliable. Tablets also show a strong approval rate of 75.2%. Consoles, smart TVs, and wearables are outliers with low transaction counts.

Figure 4.8 Distribution by device type



Finally, Figure 4.9 shows how the time taken to submit a payment form relates to approval rates, with each bar representing transaction counts in 2500ms bins. Quick submissions have a high approval rate, peaking at 85.1%, while longer times generally show a slight decline in approval. There's a plateau around 68% for submissions over 100 seconds, with minor fluctuations that are not significant. Higher approval for quick submissions might indicate familiarity or confidence, whereas lower rates for longer times could suggest repeated corrections.

Figure 4.9 Time to submit distribution. Binsize=2500ms



CHAPTER 5. RESULTS

5.1 LPM regression results

The estimation results are presented in tables below:

Table 5.1 Coefficient of LPM

Variable	Coefficient	St. dev	Statistically significant	Significant magnitude
number_autofilled_fields	-0.024	0.0003	Yes	Yes
log(time_to_submit)	0.003	0.0003	Yes	Yes
is_3ds	-0.385	0.001	Yes	Yes
is_cardholder_required	-0.129	0.001	Yes	Yes
is_postal_code_required	-0.007	0.001	Yes	No
amount_usd	-0.00001	0.00000	Yes	No
is_country_matched	0.093	0.002	Yes	Yes
is_exotic_browser	-0.046	0.002	Yes	Yes
is_exotic_cardbrand	-0.036	0.002	Yes	Yes
number_error_occurred	-0.027	0.001	Yes	Yes
number_tooltip_hovered	0.017	0.002	Yes	No
is_tracking_disabled	0.311	0.022	Yes	Yes
is_mobile	-0.038	0.004	Yes	Yes
is_desktop	0.168	0.021	Yes	Yes
is_Windows	-0.202	0.003	Yes	Yes
is_MacOs	-0.128	0.003	Yes	Yes

Table 5.1 – Continued

Variable	Coefficient	St. dev	Statistically significant	Significant magnitude
is_iOS	0.054	0.021	Yes	No
is_Android	-0.047	0.021	Yes	No
Constant	0.726	0.022	Yes	Yes
Observations	1 123 728			
R2	0.172			

The linear probability model estimated with OLS explores how different factors impact the probability of transaction approval. The dependent variable, "Approval," is binary, so using an LPM may introduce bias because it does not constrain predicted probabilities within 0 and 1, and assumes constant variance. However, the LPM provides clear, straightforward interpretations, which makes it suitable for understanding how each factor impacts approval rates.

The variable is_cardholder_required, which indicates whether cardholder information is collected on the payment form, has a coefficient of -0.129, suggesting that requiring this information reduces approval probability by about 12.9%. This might reflect the extra friction that comes with requiring additional details, which may scare off users or cause errors. Analogically, is_postal_code_required has a small but statistically significant negative effect of -0.007, indicating a slight reduction of 0.7% in approval when postal codes are required. These results highlight how certain payment form configurations that merchants can adjust have an impact on transaction outcomes by creating potential barriers.

User behavior factors also impact approval. The number_autofilled_fields variable, with a coefficient of -0.024, shows that, all things equal, each autofilled field decreases the probability of approval by 2.4%. Autofill could be associated with higher-risk transactions or users who may not double-check their entries. The log(time_to_submit) variable, with a coefficient of 0.003, slightly increases approval likelihood by 0.3%, suggesting that users who take more time to fill out the form are more deliberate or accurate, which may reduce perceived risk of error of manual input. Each additional error during form filling (number_error_occurred) decreases approval probability by 2.7%. Conversely, hovering over tooltips, number_tooltip_hovered, is associated with a 1.7% increase in approval, showing that a more careful user is expected to have higher chance of transaction approval.

Control dummy variables for user's system setup also have been estimated. The use of exotic browsers and exotic card brands, with coefficients of -0.046 and -0.036 respectively, slightly reduce approval probability, likely because less common browsers or card brands could be hard to process due to technical limitations of merchants acquiring infrastructure ot providers web form. Disabling tracking (is_tracking_disabled) significantly increases approval probability by 31.1%, which is an interesting result as it conflicts with a common assumption about tracking-enabled users being more transparent. A match between the issuing country and the IP country increases approval by 9.3%, as expected when geographic indicators align. Operating systems and devices impact approval probabilities in notable ways. Mobile users are associated with a 3.8% lower probability of approval, which could stem from increased fraud risk on mobile platforms or higher risk of mistake while entering the card data. Desktop usage, on the other hand, increases approval probability by 16.8%, perhaps reflecting greater trust in traditional desktop transactions. Among OS types, Windows is associated with a substantial 20.2% reduction in approval probability, while MacOS users also see a reduced probability, though to a lesser degree, 12.8%. iOS has a positive coefficient of 0.054, reflecting a 5.4% increase in approval likelihood, possibly linked to higher perceived stability and buying power of iOS users. Android is associated with a 4.7% decrease in approval probability.

Several transaction and card features are also influential. Transactions with 3D Secure enabled see a substantial 38.5% decrease in approval probability, possibly because this feature is often applied in higher-risk situations, where issuers are more cautious. The transaction amount in USD has a very small negative coefficient (-0.00001), suggesting that larger transaction amounts are subject to greater risk, but with a very low effect size

5.2 Probit regression results

The Probit model utilizing the maximum likelihood estimation is used to find the factors that influence the chance of the user's payment approval in online payment. Interpretation of the Probit model is quite more complicated than for the linear probability model. So we would need to find the marginal effects. Average marginal effects are presented in a Table 5.2.

Table 5.2 Marginal effects of Probit estimation

Variable	AME	SE	p-value
number_autofilled_fields	-0.0217	0.0003	<0.0001
log(time_to_submit)	0.0034	0.0003	<0.0001
is_3ds	-0.3325	0.010	<0.0001
is_cardholder_required	-0.1214	0.009	<0.0001
is_postal_code_required	-0.0038	0.009	<0.0001
amount_usd	-0.00001	0.0000	<0.0001
is_country_matched	0.0867	0.002	<0.0001
is_exotic_browser	-0.0468	0.0018	<0.0001

Table 5.2 – Continued

Variable	AME	SE	p-value
is_exotic_cardbrand	-0.0374	0.0019	<0.0001
number_error_occurred	-0.0248	0.001	<0.0001
number_tooltip_hovered	0.0200	0.0018	<0.0001
is_tracking_disabled	0.4512	0.0219	<0.0001
is_mobile	-0.0453	0.0037	<0.0001
is_desktop	0.1693	0.0211	<0.0001
is_Windows	-0.2137	0.0029	<0.0001
is_MacOs	-0.1382	0.0032	<0.0001
is_iOS	0.0540	0.0212	0.0107
is_Android	-0.0435	0.0212	0.0401

The Probit model estimation of approval probability shows average marginal effects (AME) that are generally similar to the LPM, but some differences are notable and coefficients are more trustable, because Probit model does not have limitations that are common to LPM. For instance, is_cardholder_required has a marginal effect of -0.1214, close to the LPM but indicating a strong negative effect. Requiring cardholder information reduces the probability of approval by about 12.1%, suggesting that this added form field might introduce friction or scare off users, that possibly can submit wrong information, which can indicate the riskier transaction for an issuer. is_postal_code_required also negatively impacts approval but with a smaller average marginal effect of -0.0038, indicating a small 0.38% decrease, suggesting that postal code requirements could introduce minor barriers.

Several variables show consistent effects with the LPM but differ in magnitude. is_tracking_disabled, for example, has a substantial positive effect, with an AME of 0.4512, showing an even stronger increase in approval probability compared to the LPM. This larger effect might indicate that disabled tracking is viewed as an indicator of experience of a user, because the flow of disabling tracking can be tricky. Users who know how to disable frontend tracking on the website are more likely to know how to pay online. is_3ds also has a large negative marginal effect at -0.3325, slightly smaller than in the LPM but still highlighting that transactions with 3D Secure face a higher likelihood of being declined, perhaps due to their association with high-risk transactions and some kind of sampling bias.

Other user behaviors and platform variables align closely with the LPM results. number_autofilled_fields has a marginal effect of -0.0217, suggesting a 2.17% decrease in approval probability, ceteris paribus, for each autofilled field, reinforcing the pattern that autofill is potentially related to more errors in submitting the form. is_country_matched has a positive effect of 0.0867, supporting the idea that geographical consistency between the issuing country and IP address increases approval likelihood. Device types and operating systems similarly show effects in line with the LPM, such as a positive AME of 0.1693 for is_desktop and a negative effect for is_Windows at -0.2137, both reinforcing associations with approval probability.

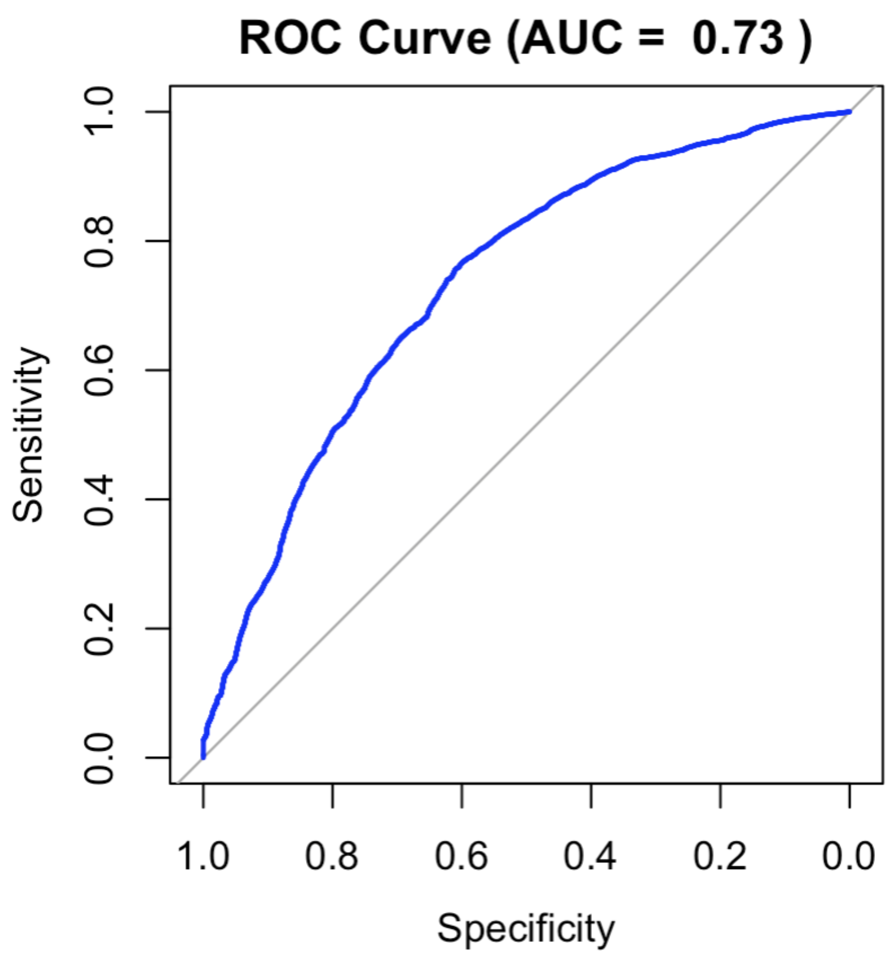
The Probit model offers similar insights to the LPM, but certain effects are changed, especcially for high-magnitude coefficients. The Probit model's results show that variables like is_cardholder_required, is_3ds, and is_tracking_disabled have a substantial influence on approval probability, with some differences in effect sizes from the linear probability model.

5.3 Model validation

To validate the model the cross-validation technique is applied. The data is randomly split up to two samples: test and train to figure out how good the model capture variations and predicts outcomes for a transactions that it was not trained on.

One characteristic for estimating the correctness of the form is receiver operation characteristic (ROC) curve. It is presented on the Figure 5.1. The area under this curve grades the model from 0 to 1. This model gets value of 0.73, which is quite good, because it moderately explains the variations in data. It would be great to have the AUC larger than 0.85, but this research and a model itself were not focused on trying to predict user's approval rate based on form configuration and user's behaviour. For the purpose of explaining variation impacted by key variables in a sample, which was the main purpose of this study, AUC = 0.73 suggests adequate results.

Figure 5.1 ROC curve, R output



Another metrics to evaluate the performance of the model is recall and precision. Precision is a measure of how often did the model correctly predict the approval of a transaction, given all predicitions about positive approval. It is equal to 0.73, which means it is not quite good at predicting. However, as I stated earlier, prediction itself is not the goal of the model. Recall is a measure of how often the model predicts true positives, given all the true positive outcomes. It equals 0.91, which is a good result for this model. This means that 91% of all approved transactions in a test sample were predicted correctly.

Also, to address multicolinearity issue, the variance inflation factor (VIF) was used while evaluating the validity of the model. This factor shows the degree to which the variables exhibit multicolinearity. A benchmark for the VIF value is 5, variables with larger VIF value were excluded from the final model. The final values of a VIF are presented in a Table 5.3

Table 5.3 VIF evaluation results, R output

Variable	VIF
number_autofilled_fields	1.113831
log(time_to_submit)	1.371940
is_3ds	1.013844
is_cardholder_required	1.178189
is_postal_code_required	1.091945
amount_usd	1.104972
is_country_matched	1.350128
is_exotic_browser	1.117344
is_exotic_cardbrand	1.081139
number_error_occurred	1.039280

number_tooltip_hovered	1.028820
is_tracking_disabled	1.304665
is_mobile	1.235085

CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

This study explored the relation between user behavior on payment forms and it's configuration on transaction approval rates in online payments. The goal was to improve understanding in this essential area of online businesses. Approval rates are important because they directly impact a business's revenue and customer satisfaction levels. The findings of this research confront some of the common assumptions about payment form configurations. Initially, we expected that requiring additional fields, like cardholder names and postal codes, would increase the likelihood of approval, as these details could add to the security of the transaction and make it seem more trustworthy. However, the results indicate the opposite — these form additional fields slightly decrease approval probability. It seems that requiring additional information from the user can introduce friction, potentially making some users to submit incorrect information which in its turn raises risk concerns on the issuers side.

Despite these surprising results, the study faced some limitations. One key gap was the lack of access to marketing data, which could have provided a more cohesive understanding of the user journey before reaching the checkout stage. For example, knowing where users came from — such as through advertisements, referral links, or directly from search engines — could clarify how their intent and trust level differ at the point of checkout. Additionally, details on what users saw during the sales funnel could help unserstand how different touchpoints impact their behavior at the payment stage. Another important limitation was the lack of knowledge about the issuer's anti-fraud logic, due to it's proprietary nature. Since issuers play a central role in transaction approvals, understanding their criteria and risk signals, which could vary significantly across regions and card types, would provide deeper context to these findings.

The research offers number of recommendations for merchants and payment providers that they can use to potentially increase their approval ratio and subsequent revenue. Merchants could simplify the payment form by minimizing non-essential fields, as our findings suggest that each additional required field slightly reduces approval rates. For instance, requiring cardholder names or postal codes should be considered optional, especially if the merchant has other security measures in place. Alternatively, merchants could employ dynamic form fields, where specific requirements appear only when needed based on the user's profile or transaction history, thus minimizing user friction.

The long-term implications of implementing these recommendations could be benefitial for both merchants and payment providers. A smoother and more user-friendly checkout process is likely to enhance customer loyalty and satisfaction, as users appreciate a hassle-free experience. Over time, merchants with higher approval rates and fewer declined transactions may see a competitive advantage, as customers are more likely to return to platforms where payments are processed easily and efficiently.

REFERENCES

Viren Rehal. 2024. Linear Probability Model (LPM): Meaning and Problems. spureconomics.com/linear-probability-model-lpm-meaning-and-problems/

Aari Dhaptel. 2024. Digital Payment Market Research Report. www.marketresearchfuture.com/reports/digital-payment-market-7572

Worldpay. Global Payments Report 2024. https://worldpay.globalpaymentsreport.com/en

Cherif, Badhib, Ammar, Alshehri, Kalkatawi and Imine. 2023 Credit card fraud detection in the era of disruptive technologies: A systematic review. *Journal of King Saud University - Computer and Information Sciences*: 145-174

Ipparagire, Barrio and Arostegui. 2023. Estimation of the ROC curve and the area under it with complex survey data. *Stat*

Rachakonda, Bhatnagar. 2021. "Extending area under the ROC curve for probabilistic labels". *Pattern Recognition Letters*: 265-271

Thompson et al. 2017. Extracting the Variance Inflation Factor and Other Multicollinearity Diagnostics from Typical Regression Results. *Basic and Applied Social Psychology:* 1-10

Senaviratna, Cooray. 2019. Diagnosing Multicollinearity of Logistic Regression Model. *Asian Journal of Probability and Statistics*

UN, 2021. COVID-19 and E-commerce. https://digitallibrary.un.org/record/3886558/files/covidecommerce.pdf?ln=en

GSMA, 2024. The mobile economy 2024. https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-economy/wp-content/uploads/2024/02/260224-The-Mobile-Economy-2024.pdf

Habibpour et al. 2023. Uncertainty-Aware Credit Card Fraud Detection Using Deep Learning. *Engineering Applications of Artificial Intelligence*

M. Krivko. 2010. A hybrid model for plastic card fraud detection systems. https://doi.org/10.1016/j.eswa.2010.02.119

A. Taha and S. Malebary. 2020. An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine. *IEEE*: 25579-25587.

Dewi, Suharman, Koeswayo, Tanzil. 2024. Factors influencing the effectiveness of credit card fraud prevention in Indonesian issuing banks. *Banks and Bank Systems*: 44-60.