MODELLING PREPAYMENT IN MORTGAGES WITH A BANK EXERCISE

A Thesis

By

Ayşin Özdil

Submitted to the

Graduate School of Sciences and Engineering
In Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in the
Department of Financial Engineering

Özyeğin University
May 2022

MODELLING PREPAYMENT IN MORTGAGES WITH A BANK EXERCISE

Approved by:

Assistant Professor Levent Güntay,
Advisor,
Department of International Finance
Özyeğin University

Assistant Professor Emrah Ahi,
Department of International Finance
Özyeğin University

Assistant Professor Gamze Öztürk Danışman,
Department of International Trade and Finance
Kadir Has University

Approval Date: 27.05.2022

To Zeki Sözen who introduced me to the world of Risk,

To my mother and colleague Burak Yüksel,

with respect and thanks…

ABSTRACT

Home loans are one of the longest-term products in the Banking sector and are exposed to multiple macroeconomic cycles throughout their maturity. Each home loan contract includes the right of the loan to pay at any time during the term of the loan which causes the risk of changes in the contractual cash flows of the Banks.

When the literature for the Turkish market is analyzed, studies are done on calculating home loan option prices and most of the calculations are based on classical option techniques. In previous studies, market spot interest rate and house price levels are used as variables. Aside from other studies in the literature, this thesis is based on prepayment probability of the home loans. Borrower-specific, loan-specific and macro-economic specific factors are chosen as the variables affecting prepayment probability. To analyze the cyclical effects, the study is carried out in one of the top 10 banks with asset sizes by selecting a 12-year data observation interval between 01.01.2010 and 31.12.2021. The study aims to model customer prepayment behavior by estimating a logistic regression using 694.778 fixed-rate home loans and 247.572 prepayment events. The data set has about 30 million observations where each home loan has monthly observations until each home loan is closed. The logistic regression model can accurately predict the prepayment behavior of contracts with an Area-under-the-Curve (AUC) statistic of 0.921 and Gini coefficient of 0.843.

In studies investigating home loans out of Turkey, interest rate, borrower income, loan to value ratio, borrower age, loan age and region are the variables that affect prepayment in the loan portfolios. This thesis shows that the most effective variables in the prepayment behaviors are the interest rate level changes, reference market interest rates, current risk of home loan and customer total debt which affect home loans payment

schedule. Interestingly, original and current Loan to Value ratio, loan maturity, loan age, customer age, customer education status and customer income have limited impact on the prepayments as indicated by the model. The most plausible for this observation is the fact that people in Turkey buy their houses for residential purposes, not for trade and do not sell their houses in a short time unless it is compulsory can be interpreted.

ÖZETÇE

Konut kredileri, Bankacılık sektörünün en uzun vadeli ürünlerinden biridir ve vadeleri boyunca birden fazla makroekonomik döngüye maruz kalmaktadır. Her konut kredisi sözleşmesi, kredinin vadesi boyunca herhangi bir zamanda ödeme hakkını içerir ve bu da Bankaların sözleşmeye bağlı nakit akışlarında değişiklik riskine neden olur.

Türkiye piyasası ile ilgili literatür çalışmaları incelendiğinde konut kredisi opsiyon fiyatlarının hesaplanmasına yönelik çalışmalar yapılmaktadır ve hesaplamaların çoğu klasik opsiyon tekniklerine dayanmaktadır. Önceki çalışmalarda değişken olarak piyasa spot faiz oranı ve konut fiyat seviyeleri kullanılmıştır. Literatürdeki diğer çalışmaların yanı sıra bu tez, konut kredilerinin erken ödeme olasılığının hesaplanmasına dayanmaktadır. Erken ödeme olasılığını etkileyen değişkenler olarak borçluya özel, krediye özel ve makro ekonomik faktörler kredi düzeyinde seçilmiştir. Döngüsel etkilerin analiz edilebilmesi için, çalışma 01.01.2010-31.12.2021 tarihleri arasında 12 yıllık veri gözlem aralığı seçilerek Türkiye'de aktif büyüklüğe sahip ilk 10 bankadan birinde gerçekleştirilmiştir. Çalışma, 694.778 sabit faizli konut kredisi ve 247.572 erken ödeme olayı kullanılarak lojistik regresyon tahmin ederek müşteri erken ödeme davranışını modellemeyi amaçlamaktadır. Veri seti, her bir konut kredisinin her biri kapatılana kadar aylık gözlemlere sahip olduğu yaklaşık 30 milyon gözleme sahiptir. Lojistik regresyon modeli, Eğri Altında Kalan Alan (AUC) istatistiği 0,921 ve Gini katsayısı 0,843 ile kredilerin erken ödeme davranışını doğru bir şekilde tahmin edebilir.

Türkiye dışında konut kredilerini araştıran çalışmalarda, kredi portföylerinde faiz oranı, borçlunun geliri, kredi teminat oranı, borçlunun yaşı, kredinin yaşı ve coğrafi bölgesi gibi değişkenler erken ödemeyi etkilemektedir. Bu tez, erken ödeme davranışlarında en etkili değişkenlerin konut kredisi ödeme planını etkileyen faiz oranı

değişimlerinin, referans piyasa kredi faiz oranlarının, ev kredisinin güncel riskinin ve müşteri toplam borcunun olduğunu göstermektedir. İlginç bir şekilde, orijinal ve güncel kredi teminat oranı, kredi vadesi, kredi yaşı, müşteri yaşı, müşteri eğitim durumu ve müşteri geliri, modelde belirtildiği gibi erken ödemeler üzerinde sınırlı etkiye sahiptir. Bu gözlem için en makul olanı, Türkiye'de insanların evlerini ticaret amaçlı değil, ikamet amaçlı aldıkları ve zorunlu olmadıkça kısa sürede satmadıkları şeklinde yorumlanabilir.

ACKNOWLEDGMENTS

I would like to thank my advisor Assistant Professor Levent Güntay who guided me with his advice and my colleague Burak Yüksel for his time, support and help during the work of the thesis and for his valuable comments. This master thesis would not have been possible without their contributions. Last but not least, I would like to thank my mother for her trust, patience and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

CHAPTER I

INTRODUCTION

Housing has been one of the first basic needs of human beings since the earliest times of history. People buy a house to meet their housing and shelter needs, as well as for different purposes especially in our new era such as for investing or trading.

Housing loans in Turkey have generally long-term maturities. According to the Financial Stability Reports issued by the Central Bank of the Republic of Turkey, their average maturity is around 100 months. In long-term loans, there may be a change in the financial strength of the borrower, as well as a change in macroeconomic conditions. And also some home loans are prepaid for rational or irrational reasons before their maturity. This thesis focuses on these longest-term products' early termination behavior and the factors affecting this behavior. In the case of Turkey in my sample data, I find that interest rate level changes and customer total debt is the most effective factors. Loan to value ratio, region, borrower age, loan age, borrower income are other significant variables found in the studies performed in the developed markets.

In the second chapter, significant milestones of the housing market in Turkey over the last 12 years are expressed. In the third chapter, risks of housing loans, deterministic and stochastic cash flows of banking products, the importance of behavioral analysis and how it changes balance sheet are described.

In the fourth chapter, detailed data of the sample and selected candidate variables are analyzed. Methodology and regression are discussed in chapter five. In chapter six, regression results are explained and interpreted.

CHAPTER II

HOUSING MARKET IN TURKEY BETWEEN 2010 AND 2021

A home loan is defined as a loan provided by a Bank (Financial Institution) or a lender (Building Contractor or Creditor) that makes it possible for an individual to own a house.

With the increasing population, housing supply and demand in Turkey and the development of financial markets, house sales (by unit) have increased more than double between 2010 and 2021. Figure 1 shows house sales between 2010 and 2021 [1].

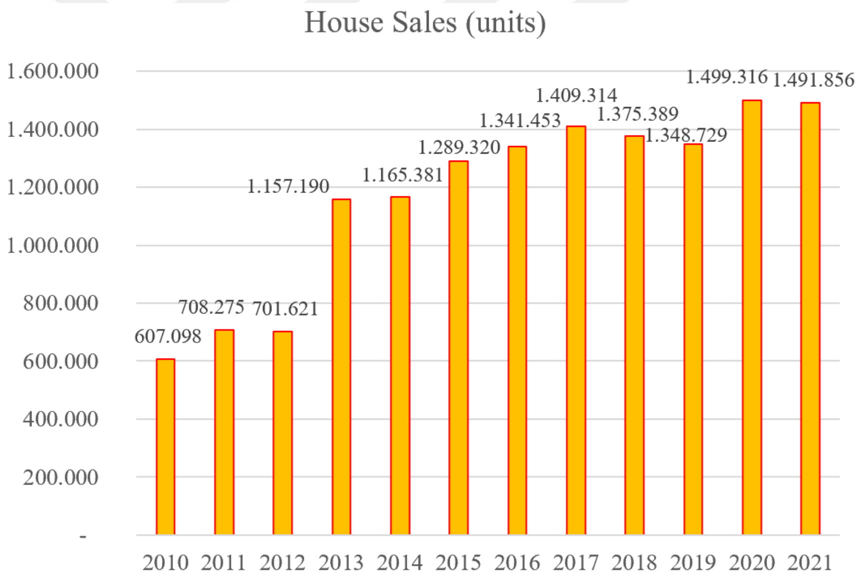

Figure 1: House Sales (units), 2010-2021

Especially in the last decade, the regulatory, agencies and the government (especially State Banks) support the housing market with legal infrastructure. The outstanding volume of home loans (in the Banking Sector) is currently over TRY 300 billion, and the volume has six times in the last 12 years. Figure 2 shows the home loan debt, the ratio of home loans and construction sector to the GDP [1][2]. The GDP calculation method is chosen by taking into account the change in market prices according to the type of economic activity. The average percentage of the home loan to GDP, 5,5%, shows the importance of the home loan debt in Turkey's economy. The main source of housing finance in Turkey is personal savings. The average saving rate of families is also around 16-20% in Turkey [3]. Although the contribution of the construction sector to GDP has been around 5% in the last period, it has been realized as 7,1% on average with a significant share in the 12 years.

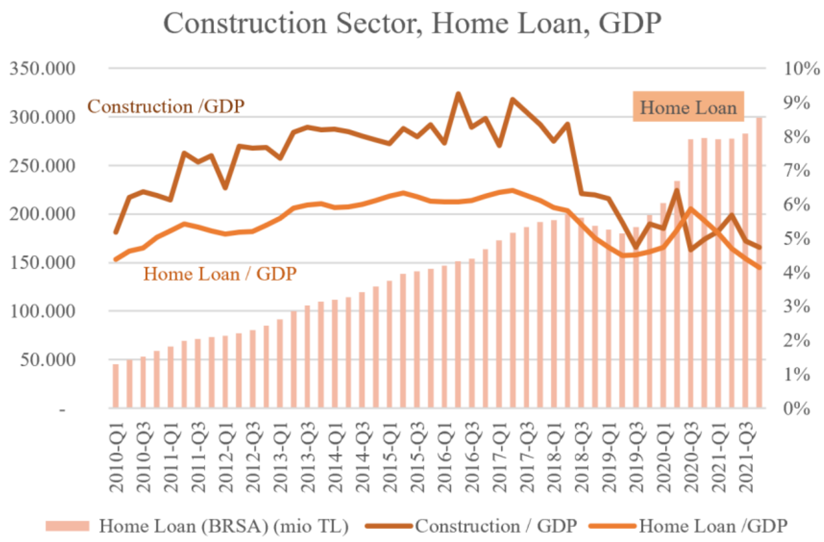

Figure 2: Home Loan and Construction Sector Vs GDP, 2010-2021

The review covers a 12-year empirical work, including only the residential market segment. With a 12-year empirical study, the research aimed to cover a full multiple interest rate cycle [2]. Figure 3 demonstrates that contract rates of housing consumer loans fluctuated between 8,3% and 28,9% between 2010 and 2021 in Turkey [1][2]. Note that the number and the share of house sales increase when the interest rate decreases. According to TURKSTAT's data, as demonstrated in Figure 1, between the period 2010-2021, the average annual house sales is 1,17 million houses [1]. Also, only 0,45 million sales (32%) are home loan sales, 68% of the sales are other house sales[1]. Usually, home

[1] According to TUIK; Mortgaged Sale: Mortgaged sales are demonstrating the same house as collateral to loan guarantee for houses purchased by borrowing. Other Sale: Including all house sales except mortgaged sales.

loans have an interest rate for a fixed period. A fixed-rate home loan is constant and will not alter for the full term of the loan. The borrower feels comfortable knowing that his/her rate or payment will not change until maturity. Long-term home loans are normally higher than short-term home loan rates [4]. In the sample data of this study, the period of home loans is about 7,5 years.

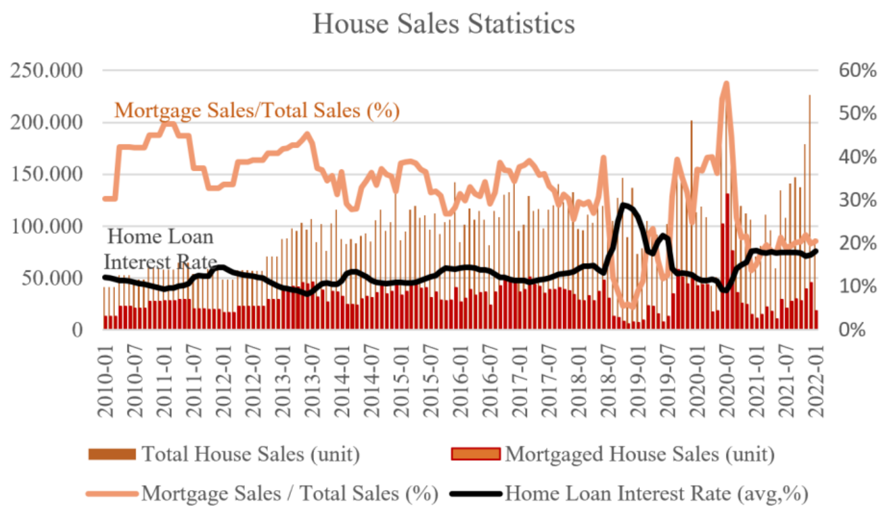

Figure 3: House Sales Statistics (unit, %), 2010-2021

In this period, the impact of foreigners buying housing due to the law amendment of Real Estate Acquisition by Foreign Persons and regular/irregular immigrants due to the civil war in Syria cannot be ignored. With the facilitating changes made in the law on owning a house in 2012, the rate of foreigners who bought a house increased rapidly [5]. In addition, there has been an increase in the number of foreigners granted residence permits due to irregular migration after the civil war. For the first quarter of the 2022, approximately 1,5 million foreigners live in Turkey with a residence permit [6]. This increase in demand has also affected the Residential Property Price Index (RPPI) on a regional basis over the years [2].

It is seen that the COVID-19 pandemic has also affected the RPPI change in the recent period. Figure 4 shows the Residential Property Price Index in Turkey (respectively; January of 2021, 2019, 2016, 2013, 2010) including the change over time at the provincial level by geographic location[2]. With the effect of migration in the eastern regions, a significant increase is noted in the index until 2019 in these regions. 2019 and beyond, with the effect of COVID-19, remote working has entered people's lives. The increase in demand for summer cottages has increased the RPPI in this southwest region.

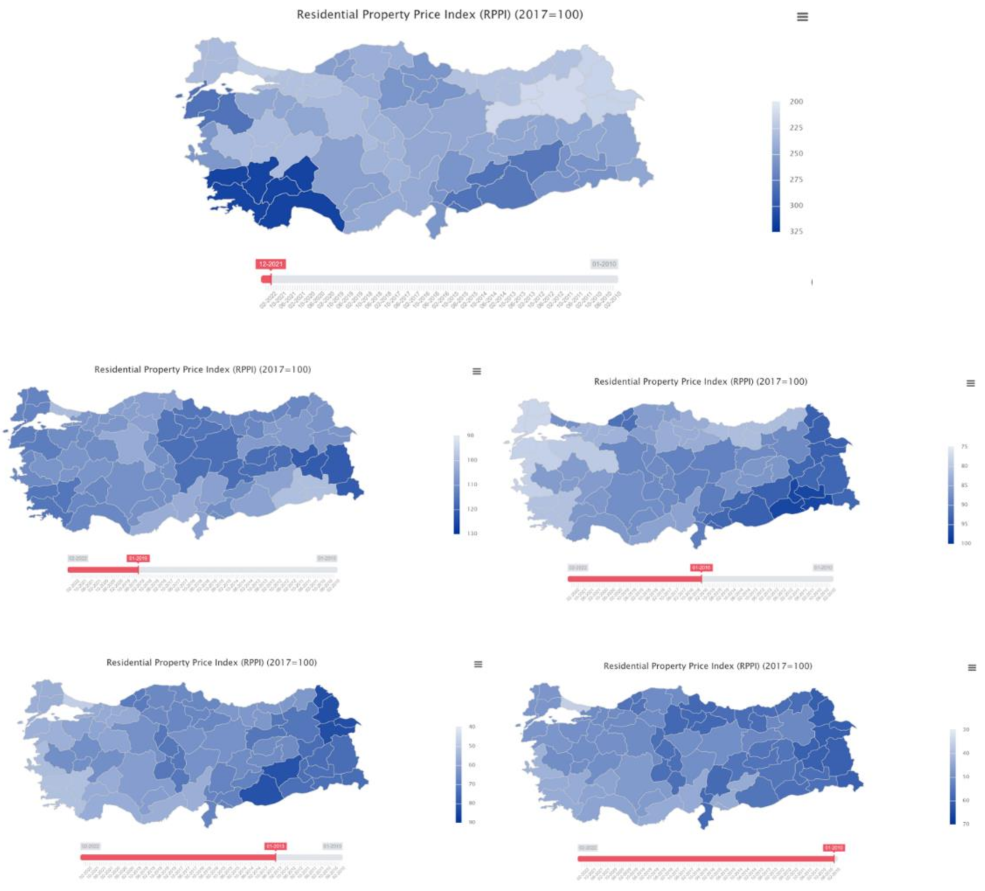

Figure 4: Residential Property Price Index (RPPI) in Turkey 2010-2021

Figures 1, 2, 3 and 4 illustrate that the Turkish House Market exhibited significant momentum to grow between 2010 and 2021.

2.1 *Home Loan Market and Regulations in Turkey between 2010-2021*

Development plans have been prepared by the Government since 1963 in Turkey. The plan aims to determine the policy to be implemented in the public sector. The plans for the last 3 periods are briefly summarized below in Table 1 in terms of the housing sector [7]–[9].

Table 1: Development Plan Policies

Development Plan	POLICY
Ninth (2007-2013)	Prioritizing disaster areas in rural settlements has been essential. It is aimed to renovate risky structures.
Tenth (2014-2018)	It is aimed to; focus on middle and lower income groups in housing production, accelerate the production of land with ready infrastructure, for home ownership strengthen the efficiency of the long-term housing finance system.
Eleventh (2019-2023)	The regulatory, supervisory, guiding and supportive role of the public in the housing market will be strengthened. By making risk prioritization of buildings in dangerous and risky areas, the country residences around and stuck in the city urban transformation services of industrial estates according to the demands and needs of the provinces will be executed.

It is understood that the housing sector had an important role in the state policy with and after the Tenth Development Plan. Housing needs have been carefully analyzed by the Government, Regulatory and Supervisory Agencies and it has been planned to be regulated in line with the results. The number of building occupancy permits accelerated in 2015-2017 in harmony with the 10th development plan as seen in Figure 5 [1]. According to the TURKSTAT, dwelling occupancy permits are decreasing due to inflation, which rose above 10% in 2018 and then rose faster, and the effects of COVID-19. Policies for the housing market are also included in the 11th Development Plan to prevent the sharp decline [9].

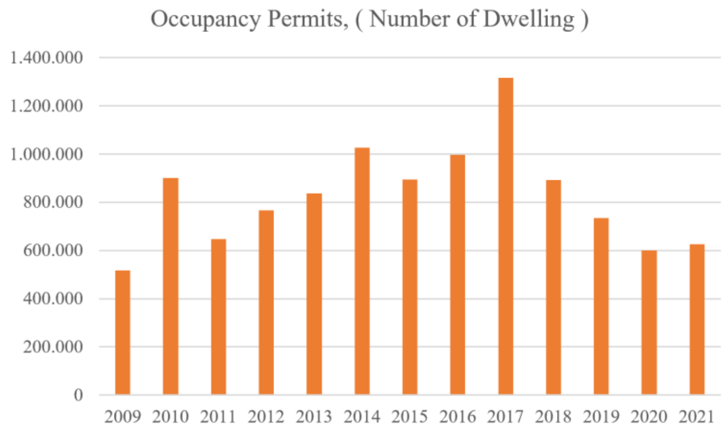

Figure 5: Occupancy Permits, (Number of Dwelling)

One of the milestones of the housing market in Turkey is the acceptance of Real Estate Acquisition by Foreign Persons (Law no.6302) in Turkey, 03.05.2012, the acquisition of the right to citizenship has been facilitated. Buying a house worth 1 million USD from 2012 to 2018 was a requirement for citizenship, this amount was reduced to 250.000 USD in 2018, and increased to 400.000 USD in 2022[10][11]. After the law amendment, according to TURKSTAT's data, approximately 295.000 houses have been sold to foreigners since 2013. The share of foreigners' house sales, which was 1,1% in 2013, increased to 4,5% in the first quarter of 2022. With this law amendment, the regular/irregular migration wave from Syria, Afghanistan, Iran, Iraq and other eastern regions, to which Ukraine and Russia are added recently, had a serious impact on the demand side of the housing market. Additionally, not only external migration but also internal migration has an impact on the housing market demand. The effect of migration from rural areas to urban in the last 10 years has also increased the demand for housing. The town and village population, which was 8,7% in 2013, decreased to 6,8% in 2021 [1]. (The law amendment, Law No. 6360.)

Since the excess demand in the housing market is more likely to increase the housing prices, the savings of those who want to buy housing will be insufficient. Due to the lack of savings, the demand for housing loans increases.

Banking rules are regulated by Banking Regulation and Supervision Agency (BRSA) in Turkey. These rules dominate the loan placement by adding restrictive or expanding rules to conditions such as loan to value, credit maturity or ratio limitations. Loan to Value ratio is one of the important ratios, which dominates the housing market. In the housing market in Turkey, it's more common to secure a loan for about 75% to 80% of the home value in 2010-2021. Table 2 shows the loan to value ratios for the years 2010-2021.

Table 2: Loan to Value Ratio Years 2010-2021

YEAR	LTV	YEAR	LTV
2010	75%	2016	80%
2011	75%	2017	80%
2012	75%	2018	80%
2013	75%	2019*	80%
2014	75%	2020**	80%
2015	75%	2021**	80%

*LTV rate cannot exceed 90% for houses with energy performance A class, and 85% for houses with energy performance class B.

** LTV rate cannot exceed 90% for houses with energy performance A class, and 85% for houses with energy performance class B. (However, for houses with a value of 500.000 TRY and below for B class energy performance LTV rate is 90%)

In the second half of the 2018, public banks started a campaign by reducing the house loan interest rates, which were in the range of 1,35%-1,25% per month, to less than 1%. This revitalized the housing market as seen in Figure 3. However, this interest rate campaign is valid for newly purchased houses and not for refinanced housing loans.

2.2 Constraints of Prepayment Home Loans in Turkey

According to the Law on Consumer Protection (6502) in Turkey, the consumer can pay one or more undue installments or pay the entire loan debt early. In such cases, the lender is obliged to deduct all required interest and other cost elements according to the amount paid early.

If the interest rate is fixed, and if one or more payments are made before the maturity of the contract, early prepayment penalty may be requested from the consumer by the housing finance institution. The prepayment penalty cannot exceed 1% of the amount (prepaid early) calculated by making the necessary interest reduction and paid early by the consumer to the housing finance institution for loans whose remaining maturity does not exceed 36 months and 2% for loans with the remaining maturity exceeding 36 months. In case the rates are determined as floating, early payment penalty cannot be requested from the consumer [12].

Table 3: Early Payment Penalty due to Maturity

Maturity	Maximum Early Payment Penalty
< 36 months	%1 x Prepaid Principal Amount
≥ 36 months	%2 x Prepaid Principal Amount

In emerging markets like Turkey with high inflation, consumers do not prefer especially long-term loans as floating interest rates. This study only focuses on fixed-rate home loans.

While many scholars and practitioners study prepayment models, most of the studies are performed outside of Turkey. Mostly based on the US and EU data. Shlomo Amar (2020) focuses on exogenous models and investigates predicting the prepayment rate (with highly nonlinear relationships between variables and prepayment events) at pool level with machine learning and artificial neural network [13]. Roberto Baccaglini, (2019[2]) investigated a logistic regression approach modeling prepayment on fixed-rate residential mortgages between September 2009 and April 2018 on a sample of a commercial bank of the UniCredit Group. Moody's Research Labs, "Mortgage Portfolio Analyzer"[3] a study by Roger M. Stein, Ashish Das, Yufeng Ding, Shirish Chinchalkar (2011) develop models to calculate prepayment probabilities over the target horizon as a function of loan-specific and economy-wide factors. Afshin Goodarzi, Ron Kohavi, Richard Harmon, Aydın Şenkut (1998) studied big-size loans in the US, how refinancing costs alter across states and countries, and which credit features make good predictor variables for prepayment forecast [14].

The study includes fixed-rate home loans and analyze the change in their payment schedule, which are prepaid between 2010 and 2021 due to Turkish macroeconomic changes. This paper covers home loans belonging to one of the top ten banks in Turkey with asset size. Loans which are defaulted and extended the maturity due to payment difficulties are excluded because they are observation data for PD (Probability of Default) models.

[2] A Guide to Behavioural Modelling, Matteo Formenti and Umberto Crespi, 2019
[3] Xufeng (Norah) Qian and Weijian Liang helped build the prepayment models.

CHAPTER III

RISKS OF HOME LOANS

The main risks for the creditor of home loan services are default risk, interest rate, liquidity and prepayment risk.

Default Risk: Borrower (individual who owned a house by credit) unable to make the scheduled payments more than three months on the debt obligation is named as default. The financial condition of the borrower may worsen because of severe illness or changing economic conditions of the borrower (unemployment, job change), making him or her fail to meet the covenants of making the monthly payment. If the house has a housing package and earthquake insurance, a credit may also be prepaid because of a fire, flood, or earthquake. If there is no insurance, the credit usually ends up in default, in case of disaster (note that it depends on laws the government can change). Also, in case of the death of the borrower, the heirs sell the house and prepay the credit. Otherwise, the loan defaults again [15]. Even though the housing loans are long-term, the average annual default rate of home loans is very low; less than 0,55 % in 2021 in Turkey's Banking Sector [16].

Interest Rate Risk: Interest rate risk includes four different types of risks. Repricing risk, option risk, yield curve risk and basis risk. Repricing risk can be determined as timing gaps in the maturity (for fixed-rate) and repricing (for floating rate) of bank on and off-balance-sheet positions [17]. Yield curve risk occurs when unexpected shifts in shape of the curve and its slope of it. Basis risk is defined as an imperfect correlation between the asset and liabilities for hedging. Option risk is a change in the cash flow on and off-balance-sheet instruments. Figure 6 shows some products includes option risk in Banking Book [18]. (For example; call options, deposit runoffs, prepayments or extensions).

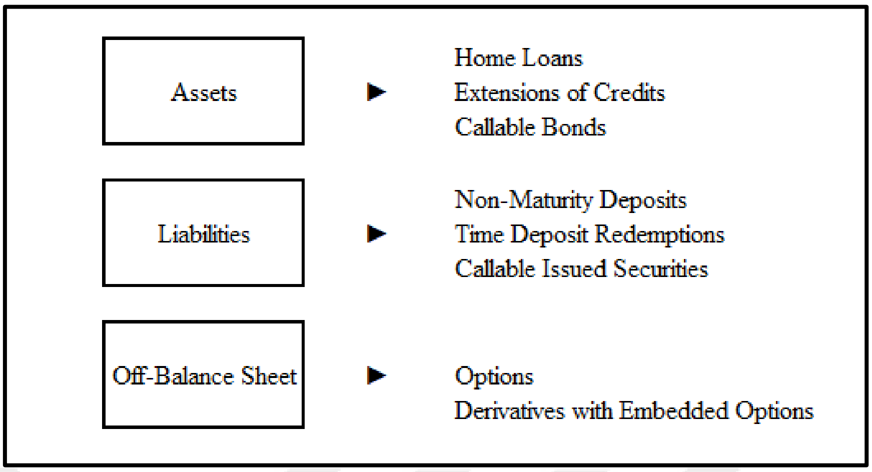

Figure 6: Some Products Includes Option Risk in Banking Book

Interest rate risk affects the bank's economic value and earnings. The traditional approach to interest rate risk is changing the earnings by decreasing the earnings or outright losses. The economic value of a Bank is calculated by the present value of net cash flows. Movements in interest rates change the financial performance of an institution. Embedded losses and gains change the bank's earnings [19]. Home loans (which include embedded options) are relatively long duration compared the other fixed-rate loans. On the customer side, the interest rate risk means the possibility of providing cheaper financing and on the Bank side, change in interest income and cash flow. Banks generally hedge these long-term home loans by using derivatives to reduce risks. The prepayment exposes banks to the risk of over-hedging, which makes the hedging in place ineffective with potential negative impacts on economic value and income statement.

Liquidity Risk: Liquidity risk includes two different risks; one of the risks is funding liquidity risk and the other is market liquidity risk. In my paper, I focus on funding liquidity risk, which means that the Bank will not be able to fund its future obligations by receiving funds from clients and other assets. The risk is that in the future the bank receives smaller than expected amounts of cash flows to fund its payment obligations. Prepayment changes the timing of the contractual principal cashflows of outstanding loans. The Gap between contractual and prepayment cash-in changes the Bank’s liquidity profile. An incorrect evaluation of this profile exposes banks to the risk of overestimating their future liquidity requirements (overfunding) as well as to the risk of increased long-term liquidity costs [20].

Prepayment is defined as only the paydown (cash) of the outstanding (unpaid) balance of a home loan pass-through that is in more of its scheduled amortization. Prepayment risk affects the financial institution from the liquidity risk perspective and interest rate margin.

The Basel Committee has decided in Interest Rate Risk and Interest Rate Risk in the Banking Book Standards[19] that banks should have to figure out the nature of prepayment risk for their portfolios and make rational and prudent estimates of the expected prepayments. Banking Book risks are associated with customers’ behavior. Many Banking products have in it implicit options such as withdrawn or repaid. To better manage the risks, Bank uses behavioral models for asset and liability balance.

The longer the duration, the greater the uncertainty and risk. Therefore, home loan is one of the riskiest products in terms of prepayment on Banking asset side. On the asset side, home loans have longer maturity compared to other assets. Also, the borrower has a right to prepay the home loan at any time during the contract. That is why home loans are called option embedded loans. The longer the duration, the more valuable the option will be. In each period until loan maturity, the borrower has an American option to use. If this option is not used, it is exchanged for a compound option to use in the future [21].

3.1 Reasons for using Behavioral Models Used in the Banking Sector

It is classified by dividing the cash flows of banks into two as deterministic and stochastic on the axis of time and amount. Antonio Castagna and Francesco Fede classify the cash flow Banks according to the probability of occurrence (as deterministic and stochastic) and the time and amount criteria. Figure 7 shows the categories that provide some examples of cash flows to which they belong [20].

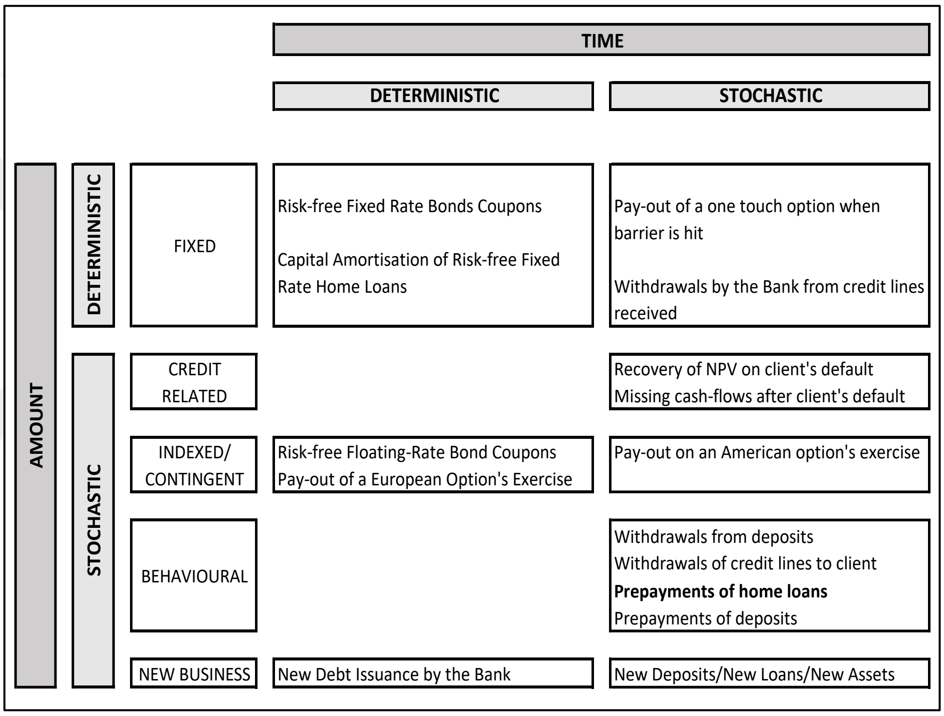

Figure 7: Classification of the Cash Flows

Deposit accounts are the main source of funding for the bank because of their low costs. Deposit contracts have no predetermined maturity, the holder is free and has the right to withdraw all or part of the amount at any time. Forecasting deposit volumes is important in every period of a macroeconomic environment because non-maturing liabilities and deposits have wide size on the balance sheet. Two main questions are important to analyze behavioral characteristics: “How much are they persistent?” and “How long do they remain stable?”

Credit cards have no commitment to a given maturity, they have revolving facilities. A practical shortcut, for uncommitted revolving products like overdrafts and credit cards, is to use an average time to closure.

Letter of credits have possible funding of irregular finance. Opening a committed credit line means the agent has the right to withdraw the amount up to the notional line whenever the agent wants under any market conditions. Credit line usage amounts are important for managing the liquidity risks and behavioral models are used for banking products to forecast the levels in some determined macroeconomic environment.

Prepayment of home loans and deposits are typical types of stochastic amounts and time, too. Banks cannot forecast the prepayment time with certainty, so the cash flows occur is stochastic, too. These uncertainties about the future have challenged the Bank managers. That forced most departments in Bank to understand the reasons for prepayment. In the most basic definition, “Behavioral Modeling” is an approach used by banks to better understand and forecast consumer actions. Economic agents (borrowers) are not taken into account standard rational principles like preferring more wealth to less wealth, sometimes for example, they can prefer to receive cash sooner than later for individual reasons. Borrowers are not fully rational and do not have access to all the information. The biggest reasons for behavioral biases in the financial markets; some economic agents have overconfidence and some show overreaction. Also for example, in some cases endowment effect, disposition effect, mental accounting, availability heuristic, certainty effect, confirmation bias, representativeness bias, regret avoidance and anchoring cause behavioral biases in the financial markets [22]. As a result of these biases, agents generally do not make optimal financial decisions. Risk management on the model side focuses on digitized or flagged variables like clients features and macroeconomic variables.

Predicting and managing cash flows is vital for banks. To manage the banking book, some products on the balance sheets of banks need to be modeled by a behavioral approach. Behavioral Models mostly used by banks are shown in Table 4.

Table 4: Behavioral Models mostly used by Banks

Interest Rate Risk and Liquidity and Funding Risk			
Prepayment/ Early Redemption Models	Volume Attrition Models	Rate Stickiness Models	Draw Down Models
General Purpose Loans Commercial Loans Home Loans Term Deposits	Sight and Term Deposits Credit Cards Overdrafts	Sight and Term Deposits Administrative Rate Loans	Letters of Credit Guarantees Committed Credit Lines

In summary, banks spend time on behavioral model studies to avoid the risks outlined below [23];

- Reduce interest rate margin
- Negative impacts on economic value
- Overestimate funding cost
- Overhedge its loan portfolio
- Alters the timing of the expected cashflows

3.2 The determinants of Home Loan Prepayments

Home loans are commonly paid off before their termination date. As stated before, the possibility of prepayment increases due to the long duration of home loans. The literature for residential home loans on prepayment models can be categorized into two groups: Optimal and Exogenous Prepayment [24].

Optimal prepayments assume that prepayment is exercised rationally. The borrower prepays when the value of the home loan is greater than sum of the present value of the outstanding debt and the prepayment costs. Van Bussel (1998) theorized optimal prepayments and calculated home loans as callable bonds.

Exogenous prepayments assume that realized prepayments are often exercised in irrational way or non-optimal. In other words, the borrower prepays when the value of the home loan is below the outstanding debt and the prepayment costs. Prepayments contain important stochastic variables following behavioral uncertainty. Exogenous models use different explanatory variables. These models aim to determine the observed prepayment rates with a set of explanatory variables.

Table 5: Exogenous Models

Exogenous Models	»	Dunn and Mc Connell (1985) Brennan and Schwartz (1985) Johnston and Van Drunen (1988) Richard (1991) Kau and associates (1992) Gliberto and Ling (1992) Archer and Ling (1993) Stanton (1995) J.M. Quigley, Y. Deng, and R. Van Order (2000)

Popular statistical techniques for exogenous modeling prepayments are Cox Proportional Hazard Models (Survival Analysis) and Logistic Regression (Binary Choice Models). Survival analysis focuses on calculating the expected time until prepayment events occur and the probability distribution for the duration of the home loan. Logistic regression is popularly used by banks to forecast whether home loans will prepay or not.

Prepayment[4] occurs because of two main categories: Financial and Non-Financial Reasons.

<u>Financial Reasons (Macroeconomic Reasons)</u>[25]: The reasons for prepaying a mortgage may depend on interest rate paths (refinancing in a declining interest rate

4 Prepayment means full prepayment of the all outstanding loan obligation in this thesis, I do not analyze partial prepayments in this study.

environment), factors like central banks' weighted funding costs, housing prices (market prices), employment status or negative income shocks/changing job or unemployment resulting in a move. Default, foreclosure or sale.

Non-Financial Reasons (Behavioral Reasons): Because of health or family problems, divorce or a different involuntary situation like foreclosure or sale, or willingness to sell to move a bigger house (family size), disasters (fire, flood or earthquake).

Factors that influence the borrowers' decisions can be divided also into three parts: Borrower-specific factors, loan-specific factors and macro-economic factors.[26]

Borrower-specific factors: Borrower age, loan purpose, the total debt of the borrower, borrower income, time with bank, borrower education status, etc.

Loan-specific factors: Loan age, loan term, loan amount, home loan interest rate, market value of the property, original value of the property, penalty, etc.

Macro-economic factors: Housing prices (RPPI), home loan interest rates, risk-free rates, benchmark 2y and 10y interest government bond yield, divorce rate, etc.

Prepayment may come with full prepayment or overpayment (partial prepayment). This paper contains "full prepayment". Variables containing the above-mentioned causes and factors are selected for the model.

CHAPTER IV

DESCRIPTION OF THE DATA-SET

4.1 Data Structure

The home loan data sample for this paper belongs to one of the biggest top ten Banks in Turkey. The database consists of 694.778 housing loans. Fixed-rate retail home loans from the portfolio between January 2010 and December 2021 are selected. Not only floating rates but also defaulted home loans are excluded. There were only 506 floating rates and 9.585 loans defaulted home loans. Loans that extend the maturity due to financial difficulties are also excluded. There are 878 restructured loans due to the payment difficulties and not in the data.

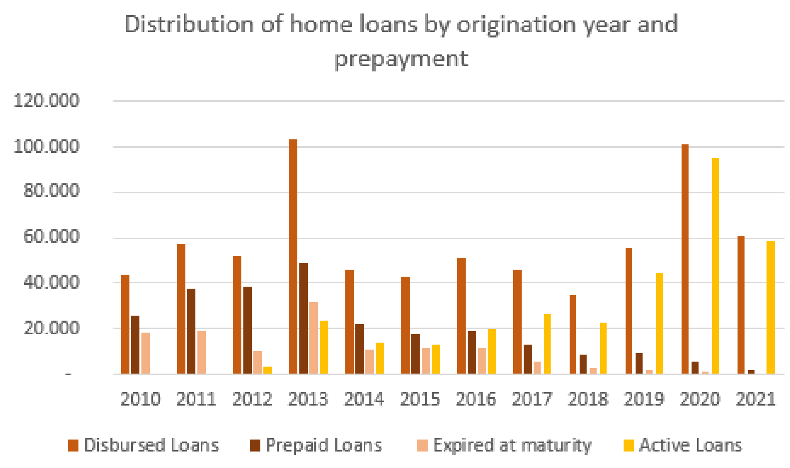

Figure 8: Distribution of Home Loans by Origination Year and Prepayment (Prepayment Years are 2010-2021)

As visible in Figure 8, the bulk of the active portfolio originated in the 2019-2021 period. Approximately 60% of the loans disbursed between 2010 and 2013 were prepaid in the following periods. Out of 43.595 credits disbursed in 2010; 58% prepaid, 41% pay on time, only 56 credits are active. Similarly, approximately 70% of the loans extended in 2011 and 2012 are prepaid. After 2013, with the upward movement in market interest rates, prepayment rates also decreased. It is not surprising after sharply decreasing interest rates in 2013, as shown in Figure 9 [2]. 2013 and 2019 are active years in housing loans for disbursing home loans. This is in line with interest rates.

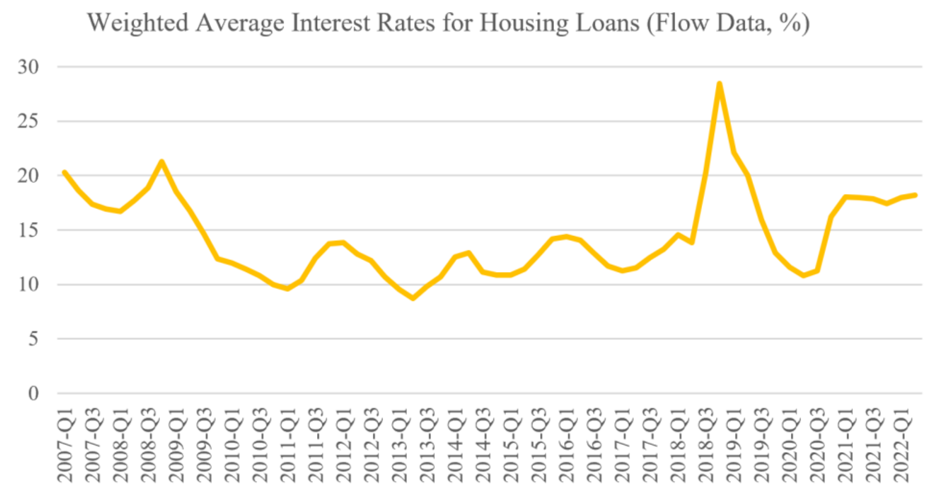

Figure 9: Weighted Average Interest Rates for Banks Housing Loans (Flow Data, %)

As can be seen in Figure 8 and Figure 9 the interest rate cycle between 2010 and 2014 may have caused prepayment. In addition, 2013, the lowest point of interest rate of housing loan interest rates, shows the maximum level of the housing loan disbursements. 15-year time period includes multiple cycles, I can explain higher prepayment rates in 2010 and 2011 with the sharp decline after 2008. In housing loans that are disbursed in 2019, prepayment in those housing loans in 2020 and 2021 are not existed because of the sharp decline in housing interest rates. Interest rates in this period are valid for new housing loans not for refinancing housing loans. Also, in this period, the homeowners do not sell their houses for trade and do not buy a new one (or negligible for my portfolio for 2020 and 2021).

It is seen in Figure 8 that 60% of the disbursed loans in 2010-2013 are prepaid. The loans disbursed in the last 5 years constitute 77% of the active portfolio.

In Table 7, the disbursement year and prepayment year of the loans are detailed. Due to the rising housing loan rates after the second quarter of 2018, it is not rational to prepay in 2018 and 2019. But, in this period, some home loans are prepaid. Prepaid loans age in 2018 has been 47 months and in 2019 has been 53 months, as shown in Table 7. Prepaid loans in 2018 and 2019 are older than those disbursed between 2010 and 2013 (which prepaid the following years). Table 8 shows that approximately 40 months is the average age for prepaid loans disbursed between 2010 and 2013.

Table 6: Prepayment Distribution by Year

Disbursed Year	Disbursed Unit	Prepaid Unit	Prepayment %	Prepayment Percentage by Year												
				2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022*
2010	43.595	25.448	58%	2,2%	5,3%	5,5%	17,9%	7,0%	6,8%	3,6%	3,3%	2,0%	2,4%	2,3%	0,0%	0,0%
2011	57.309	38.063	66%		1,1%	9,5%	29,5%	4,7%	5,2%	4,7%	3,1%	2,1%	2,2%	3,5%	0,9%	0,0%
2012	51.998	38.376	74%			10,9%	39,6%	3,5%	4,5%	3,8%	3,9%	1,6%	1,8%	2,3%	1,6%	0,2%
2013	103.746	48.723	47%				3,6%	4,0%	5,8%	6,6%	7,0%	5,1%	4,5%	6,4%	3,7%	0,4%
2014	46.244	22.028	48%					5,7%	8,7%	6,0%	6,9%	4,9%	5,5%	5,9%	3,6%	0,4%
2015	42.746	17.457	41%						1,7%	5,9%	7,9%	5,7%	6,3%	8,7%	4,1%	0,5%
2016	50.908	18.844	37%							2,4%	6,7%	4,8%	6,1%	10,5%	6,0%	0,5%
2017	45.915	13.317	29%								1,4%	3,5%	5,2%	11,0%	7,1%	0,8%
2018	34.547	8.965	26%									1,5%	5,8%	10,9%	6,9%	0,8%
2019	55.602	9.184	17%										1,6%	7,7%	6,4%	0,8%
2020	101.500	5.432	5%											1,1%	3,7%	0,5%
2021	60.668	1.735	3%												2,3%	0,6%
SUM	**694.778**	**247.572**	36%	0,1%	0,4%	1,9%	7,1%	2,1%	2,7%	2,8%	3,3%	2,5%	3,1%	5,3%	3,7%	0,5%

*First 3 months

Table 7: Age of Prepaid Home Loans by Disbursed and Prepayment Year

Disbursed Year	Prepaid Loans' age by Prepayment Year												
	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022*
2010	6	10	24	34	47	57	70	81	93	107	115	125	144
2011		4	13	21	37	49	60	73	85	99	110	117	124
2012			7	8	22	32	45	56	67	81	91	105	111
2013				4	14	25	37	49	59	73	84	97	104
2014					6	11	24	35	47	60	71	84	90
2015						4	13	23	35	49	59	72	80
2016							5	12	23	37	47	59	67
2017								4	10	24	34	47	54
2018									4	14	25	38	45
2019										4	10	23	29
2020											4	13	19
2021												4	6
Loan Age when prepaid (month)	6	9	12	16	25	31	37	41	47	53	54	52	52
Loan Term if not Prepaid (month)	91	89	98	100	83	82	80	83	83	88	93	97	99
Borrowers Age who prepaid	40	41	40	40	41	42	42	42	43	43	43	43	44

Table 8 represents that the contractual loan term of the home loans (original) and (behavioral) prepaid home loans are the same as 94 and 91 months, nearly the same only 3 months difference (Age is weighted by loan disbursement amount). It is hard to say that long-term or short-term home loans are more prepaid than others in our portfolio. The age of prepaid loans is 34 months, which is strictly far away from the average contractual loan term. Prepaid and expired-at-maturity home loans have only 9 months contractual loan term difference on average. Table 8 calculates the loan age based on the year of disbursed. Therefore, it produces more meaningful results for the years with a sufficient observation period. For example, for 2010 disbursed loans, it calculates the age of loans disbursed in 2010 and prepaid early on any date. 12 years is a sufficient observation period to calculate the loan age. In contrast, loans disbursed in 2020 can only be prepaid in 2020 and 2021, so the age of loans prepaid in 2020 remains relatively young. Sufficient observation time does not apply to these loans.

Table 8: Distribution of Home Loans by Origination and Loan Age in Months

Loans			Passive Loans (Expired)			Loan Term of Active Loans
Year	Unit	Loan Term of Origination Date	Loan Term of Prepaid Loans	Age of Prepaid Loans	Loan Term of Expired at maturity	
2010	43.595	88	92	48	60	178
2011	57.309	99	102	39	55	172
2012	51.998	92	95	23	31	120
2013	103.746	89	88	50	58	119
2014	46.244	90	90	39	45	117
2015	42.746	83	81	40	45	113
2016	50.908	85	82	36	42	110
2017	45.915	88	82	32	35	102
2018	34.547	87	81	25	28	97
2019	55.602	93	85	15	18	98
2020	101.500	117	108	11	13	118
2021	60.668	91	93	4	5	91
Sum	694.778					
Average Age (month)		**94**	**91**	**30**	**36**	**107**

So far, loan disbursements and prepayments have been analyzed on a yearly basis. Similarly, loan disbursements and prepayments are analyzed the based on geographical location and province basis as follows.

As seen in Figure 10, Figure 11 and Figure 12 that the loan disbursements of the big provinces and the related prepayments are numerically higher in the same direction. However, when prepayments and loan disbursement are proportioned on a provincial basis, there is no situation that distinguishes big cities in terms of financial literacy.

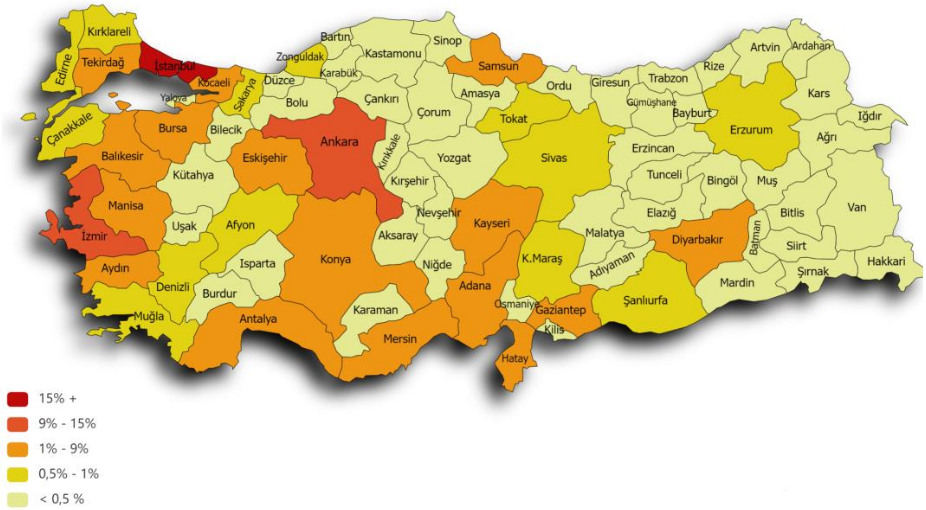

Figure 10: Distribution of Home Loan Disbursement Percentage According to Geographical Location

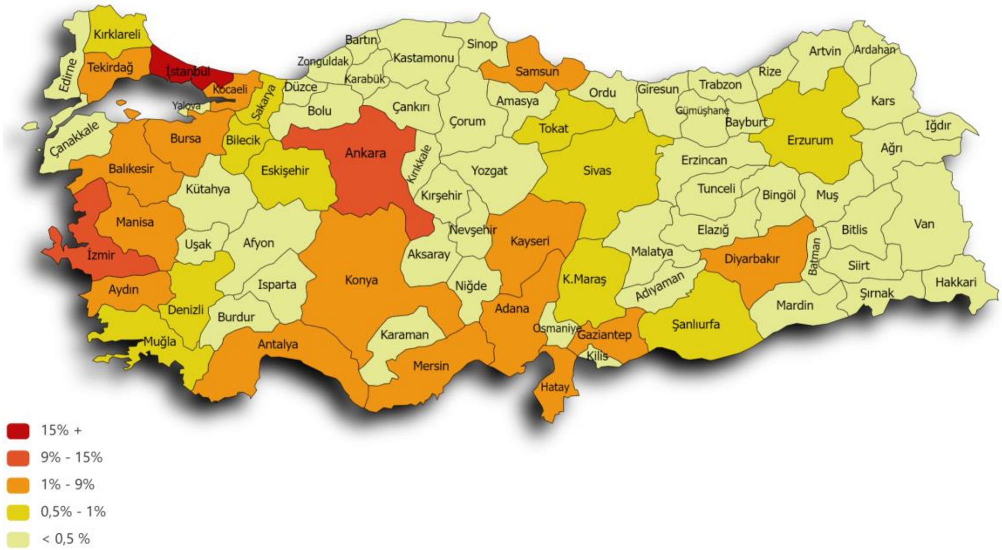

Figure 11: Distribution of Home Loan Prepayment Percentage According to Geographical Location

It is not interesting that almost half of the home loans originated in three big cities (İstanbul 24,7%, Ankara 15%, İzmir %9,65) in Turkey. The maps in Figure 10 and Figure 11 show the disbursements and prepayment percentages by province. The more home loan disbursement the more prepayment occurs as shown by Figures 10 and 11.

In Figure 12, the ratio of Prepaid/Disbursed home loans for each city is colored and shown by province. Interestingly that neither high house price index cities nor big cities have the highest ratio. In the next chapters will be clearer that neither customer education nor house price index is significant variables in our model. Çankırı (47%), Ardahan (42%), Kırşehir (41%), Bayburt (40%) are the provinces with the highest prepayment of loans according to prepayment to loan disbursement ratio in the city based. Although these provinces have high prepayment rates due to their small number of loan disbursements, the occupational groups of the prepaying customers are examined. Note that the customers in the range of 35%-45% are from occupational groups such as soldiers, teachers, doctors and polices. I can get the impression that these people, who are constantly working in different places for professional reasons, have a high house turnover rate. The geographical map results are in line with the model results.

Figure 12: Distribution of Prepaid/Disbursed Percentage for Each City

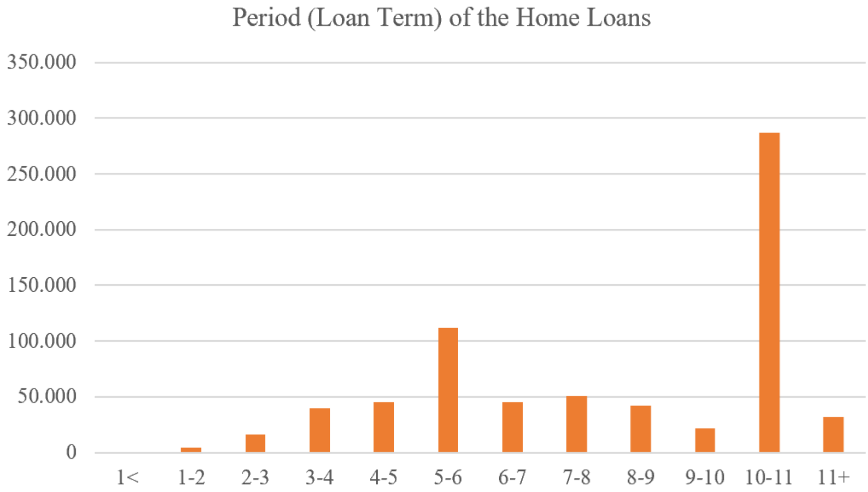

Figure 13: Period (Loan Term) of the Home Loans

In Figure 13 periods of disbursed home loans are shown. Home loans with a period of 120 months are the most common term among the data, which represents the 41% of the portfolio. Moreover, 60 and 71 months are also popular. The original term-to-maturity of the home loan did not turn out to be significant in the model.

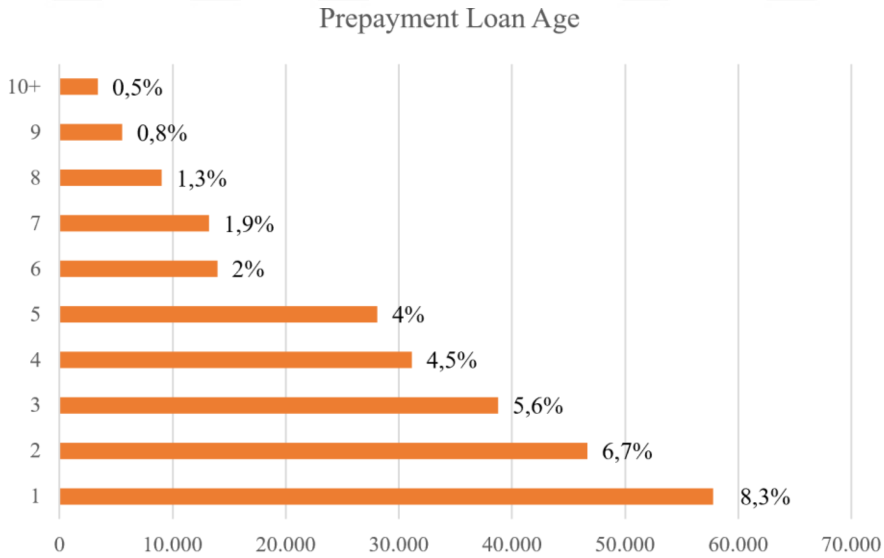

Figure 14: Prepaid Home Loans by Loan Age

In Figure 14 frequencies of prepaid loans are shown with their loan ages. 8,3% of the disbursed loans are prepaid within 1 year from the date of disbursement in our 12-

year period. 20% of the disbursed loans are prepaid in the first three years after the first installment date.

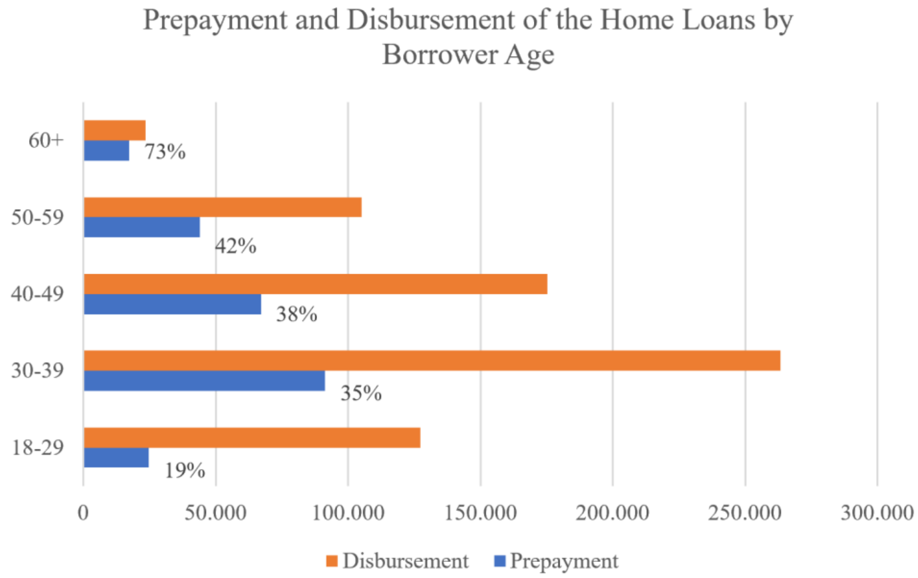

Figure 15: Prepayment of the Home Loans by Borrower Age

The age of the borrower may be a significant variable that may differ in behavior. This is logical because, elderly people seem to be less related to the market moves. They are reluctant to move. Different age cohorts of the borrowers in Figure 15 show the prepayment percentages, prepayment and disbursement units. As the age of the customer increases, the number of prepayments decreases in units, but the percentage of prepayment increases. Young customers do not have sufficient cash flow to prepay the home loan in Turkey can be inferenced. It is not surprising, in our model, that the age of the borrower has insignificant impact.

4.2 Variable Selection for the Model

The prepayment attitude at the loan level is tried to explain with three different types of variables: Borrower-specific, loan-specific, and macro-economic factors.

Borrower and loan-specific factors are the intrinsic attributes of the contract, which the Bank calculates and dominates the loan specifics (like amount, term, etc.)

according to the borrower's eligibility. Borrower's eligibility is determined by various factors like Turkish Credit Registration Bureau (KKB) score or Banks internal score, disposable surplus income, family members or spouse's income, documented income (like rent earned). The Bank also takes into account the age of the borrower. The borrower's age and maturity of the loan are important because above a certain age is risky, most banks define a threshold for the age of the borrower. By law, banks expect borrowers to make a 20-25% down payment. Value of the property, loan term, down payment amount, loan amount, monthly loan payments are other important factors. The whole point of the Bank is to make sure borrowers have the capacity to repay the home loan.

The loan-level data is completed with macro-economic factors, which may affect home loan prepayment decisions. Macro-economic factors representing the economic conditions of the country are selected. Table 11 and Table 12 show the abstract of the selected macro-economic factors. I provide the macroeconomic data for Turkey from Electronic Data Delivery System (EVDS), Turkish Statistical Institute (TURKSTAT) and Banking Regulation and Supervision Agency (BRSA) web sites.

Derived variables are available for both borrower-specific, loan-specific and macro-economic factors to show the time effect. Selected candidate variable lists are created in Table 9 below.

Candidate variables described below is a data table that contains snapshots of the home loan perimeters at the single contract level and a fixed monthly frequency.

Table 9: Categorizing Selected Candidate Variables

Selected Candidate Variables	
Borrower and loan specific factor variables	Macro-economic factor variables
Derived borrower and loan specific variables	Derived macro-economic factor variables
Derived loan and macro-economic specific variables (interest ratio/spread, PVR,..)	

Selected Candidate Variable List Borrower and Loan-specific factor variables:

Score is one of the most important variables for borrower-specific factor. But, the credit score is not available during the first two years of the database between 2010 and 2011. Furthermore, credit scores are obtained from the Turkish Credit Registration Bureau occasionally whenever the customer credit condition is investigated and are not updated regularly. For these two reasons, to prevent the selection bias I do not select the credit score as an independent variable in the model. Also, income is another important variable but have minor effect as an explained variable. Therefore, other income indicator variables such as salary flag in the bank and customer education status are used. The length of time the borrower first used a loan at the bank, which is an indicator of financial literacy and credibility, are also added to the selected candidate list as a variable. In order to avoid inflation and unstationary effect, variables expressed by Turkish Lira like total risk, borrower total debt, current principal balance, customer income are divided by the inflation rate and its log are taken.

Table 10: Selected Borrower and Loan Specific Candidate Factor Variables

VARIABLE	DESCRIPTION
Loan Term	The number of months the borrower will be making payments towards the home loan
Current Principal Balance [log (Current Principal Balance /CPI)]	The amount of principal outstanding of home loan for each date of observation
Borrower Total Debt [log (Borrower Total Debt/CPI)]	Total debt of the customer in the bank based on customer (not only home loan debt, full debt of the customer in the Bank)
Total Risk [log(Total Risk/CPI)]	The amount of principal outstanding, interest, insurance of home loan for each date of observation
Customer Age	Customer age for each date of observation
Interest Rate	Contract interest rate on a monthly basis
First Loan Time	The time between active first loan disbursement and reference date (activity period in Bank)
Time with Bank	The time between giving the customer id and reference date (time of meeting the customer with the bank)
Original Value	The purchase price of the property on the credit disbursed date
Original LTV	Original principal debt divided by the purchase price of the property
Customer Education	Customer education status (ordinal order)
Loan Age	Observation date minus Disbursement date is defined as loan age monthly basis
Salary Customer Flag	Customer receiving salary from bank is defined 1, other 0
Customer Income [log(Customer Income/CPI)]	Customer's income/salary on a monthly basis

Selected Macro-Economic Candidate Factor Variable List:

Economic conditions have a significant impact on prepayment rates. A weakened housing market, high unemployment and interest rates create recession for home loan borrowers and decrease the incentive to prepay. Changing liquidity costs affect the home loan market.

The Residential Property Price Index (RPPI) is vital to calculate the current loan to value ratio. Thanks to the payments made by the borrower until the observation date, the borrower's debt is reduced, and the current value of the house is calculated by the RPPI index.

Turkey's economy is sensitive to FX rates because of this not only interest rates but also FX levels, monthly and yearly change of FX levels (%) are selected macro-economic factors. The variable USD/TRY level which is in a constantly increasing trend, is divided by the inflation rate and its log is taken. By this way the variable which is continuous trend, is taken into account as a stationary variable.

Housing loan weighted average interest rates published by the CBRT are taken as the reference interest rate that is subject to prepay housing loans.

Inflation data, which is important for emerging markets, is also included in the selected candidate list.

To show the movement of market and housing market, Sector Housing Loan Amount, Total Industry-Level, Capacity Utilization Rate of Manufacturing Industry and Benchmark Government Yields are selected, too. Similar to USD/TRY level variable, Sector Housing Loan Amount is in a constantly increasing trend. Same method[5] is applied to detrend the variable.

[5] Log (Variable expressed as level for each month/CPI of the same month)

Sometimes unexpected customer behaviors may be encountered due to family problems. That is why divorce rates have also been chosen as a variable. Divorce rates are published annually by TURKSTAT. Therefore, it can not be included as a significant variable in our model.

Table 11: Selected Macro-economic Candidate Factor Variables

VARIABLE	DESCRIPTION
Residential Property Price Index (RPPI) - Level and monthly change (%)	Observable characteristics of the houses represent the whole of Turkey. The hedonic regression method was used to track price changes adjusted for quality-related effects.
Reference Interest Rate	Weighted average interest rates for banks' housing loans (Flow Data, monthly %)
CPI (Consumer Price Index)	It refers to the Consumer Price Index (y-y,%) published by the TURKSTAT monthly.
Sector Housing Loan Amount monthly change (%) [log (Sector Housing Loan Amount Level / CPI)]	It refers to the housing loan amount included in the weekly sector consumer loan published by the BRSA.
USD/TRY monthly and yearly change (%) [log (USD/TRY Level/CPI)]	Expresses the foreign exchange sales level published by the CBRT
Unemployment Rate	The ratio of the unemployed population in the labor force. Unemployed means in TURKSTAT's definition.
Total Industry-Level	It refers to the Industrial Production Index published by the CBRT (TP SANAYREV4 Y1) and TURKSTAT.
TRY2Y-end	Benchmark 2y interest government bond yield, the source is Reuters.
TRY10Y-end	Benchmark 10y interest government bond yield, the source is Reuters.
Capacity Utilization Rate of Manufacturing Industry	It refers to the Capacity Utilization Rate of Manufacturing Industry-Level published by the CBRT (TP.KKO2.IS.TOP) and TURKSTAT.
Divorce Rate Level and annual change	It refers to the number of divorces per 1000 population in a given year.

Derived borrower and loan-specific from selected candidate variables:

I have derived two variables by relating the borrower's specific factor with loan-specific factor. One of them is the ratio of the penalty amount to the borrower's income. The penalty amount is calculated with the remaining maturity and current principal balance for each date of observation. If the remaining maturity is less than 36 months, then the penalty amount is equal to the 1% of the current principal balance, if it is greater than the 36 months then equal to the 2% of the current principal balance.

The other variable is loan age that the period between the loan disbursement date and the reference date in months. It will show us how long the loan has lived and how old it is when it closes early.

Derived macro-economic factors from selected candidate variables:

I derived new macro-economic variables by taking the differences in some of the macroeconomic data listed above. Banking sector consumer housing loans (year to year percentage change), USD/TRY (year to year percentage change), Residential Property Price Index (RPPI) (year to year percentage change). Also, I added the share of housing loans in total consumer loans as a percentage change, too.

Lastly, I add "Yield Curve Slope" (Litterman, Scheinkman & Weiss, 1991)[14]. The difference between the benchmark government bonds TRY10Y and TRY2Y. I calculate the new slopes for each date of reference to analyze the effects of short and long term benchmark interest rate changes.

$$\text{Yield Curve Slope} = \frac{\text{TRY10Y} - \text{TRY2Y}}{\textit{Monthly Average}\ (\text{TRY10Y} - \text{TRY2Y})} \qquad (1)$$

Derived loan and macro-economic factors from selected candidate variables:

Some loan-specific variables changes by the macro-economic conditions. Changing the interest rate risk environment can make a house loan cheaper or more expensive. To calculate the time effect, I derive new variables listed below in Table 12.

In order to avoid inflation and unstationary effect, variables expressed by Turkish Lira like current penalty is divided by CPI and its log is taken.

Table 12: Derived Loan and Macro-economic Specific Variables

VARIABLE	DESCRIPTION
Interest spread	Home Loan Contract Rate minus Current Refinance Rate
Interest ratio	Home Loan Contract Rate divided by Current Refinance Rate
Present Value Ratio	PVR=I/R*[1-(1+R)^(-M)]/[1-(1+I)^(-M)]
Prepayment Option Value	(NPV(Remaining Home Loan Debt)-(Cost of Closing Home Loan with Penalty))/(Cost of Closing Home Loan with Penalty)
Current Value	(Current) Market value of the property for each date of observation
Current LTV	Current (remaining) principal divided by the market value price of the property for each date of observation
Current Penalty [log(Current Penalty/CPI)]	The current penalty amount calculated according to Table 3 for each observation date
Current Penalty to Current Value Ratio	The ratio of the penalty amount to the current property value for each date of observation
Current Penalty to Customer Income Ratio	The ratio of the penalty amount to the customer income for each date of observation

The interest rate is the most important driver in rational calculations for prepayment. Interest rate has linear and non-linear effects on prepayment that is why I derive the “interest spread” and “interest ratio” for each date of observation’s reference interest rates.

$$\text{Interest Spread} = \text{Home Loan Contract Rate} - \text{Current Refinance Rate}$$

$$\text{Interest Ratio} = \frac{\text{Home Loan Contract Rate}}{\text{Current Refinance Rate}} \qquad (2)$$

I also add the Present Value Ratio (PVR) (Richard & Roll, 1989). This ratio includes the current (reference) market home loan interest rates and the contract interest rate also includes remaining term of the loan.

$$\text{PVR} = \frac{\text{I}}{\text{R}} * \frac{[1 - (1 + \text{R})^{-M}]}{[1 - (1 + \text{I})^{-M}]} \qquad (3)$$

I= Contract Rate on a monthly basis

R= Current Home Loan Refinance Rate on a monthly basis

M= Remaining Life of the loan in months

Another calculated and added variable is 'Prepayment Option Value' for each loan and for each observation cohort date.

$$\text{Prepayment Option Value} = \frac{\text{NPV(Remaining H. Loan Debt)} - (\text{Cost of Closing H. Loan with Penalty})}{\text{Cost of Closing H. Loan with Penalty}} \quad (4)$$

$$\text{NPV(Remaining H. Loan Debt)} = \frac{\text{Installment}}{R} * \left(1 - \frac{1}{(1+R)^{M}}\right) \quad (5)$$

R= Current Home Loan Refinance Rate on a monthly basis

M= Remaining Life of the loan in months

Installment= The amount that every month paid by borrower

$$\text{Cost of Closing H. Loan with Penalty} = (\text{Penalty} + \text{Cur. Principal Balance}) * 1{,}02 \quad (6)$$

The current value of the house is calculated by the province where the house is located and the housing price index of the province for each date of observation. Similarly, the current loan to value ratio is calculated by the current (remaining) principal divided by the market value price of the property for each date of observation.

Additionally, "Current Penalty to Current Value Ratio" is calculated for each date of observation, too.

How do I define and flag the prepayments?

In data checks, I observe that long-term housing loans with a maturity of 120 months, for example, is closed only 1-2 months or 3-5 days ago.

And such long-term loans, I analyzed that prepaid relatively short time before the maturity should not be flagged as prepayment. I have defined one more new variable (ratio) for prepayment flagging that I mentioned below:

$$\text{Unused Maturity Percentage} = \frac{\text{Loan Term} - \text{Loan Age on Prepayment Day}}{\text{Loan Term}} \quad (7)$$

Table 13: Derived Variables for Prepayment Flag Study

VARIABLE	DESCRIPTION
Unused maturity Percentage	Difference between home loan term (as a month) and loan age (as a month) on prepayment day divided by loan term (as a month)

Inspired by the 90% confidence level, if the loan status code is closed and then if the unused maturity percentage ≤ 10% then I flagged it as a prepaid loan. I also have calculated for the 5% level and found the difference between the prepayment number as 23.593, (equals to 3,3% of the portfolio), for the two levels. For a clear understanding, I have explained in Table 13 and give examples in Table 14.

Table 14: Prepayment Examples

Loan Term	Age at Closed	Unused Maturity Percentage	Prepayment
46	26	43,5%	1-Yes
113	25	78%	1-Yes
120	107	11%	1-Yes
120	108	9%	0-No
40	36	9%	0-No
100	93	7%	0-No
60	57	5%	0-No
120	118	2%	0-No
90	90	0%	0-No

The home loan data covers the period between January 2000 and December 2021. The database contains snapshots of home loan parameters and macroeconomic variables with monthly frequency. I have selected this period because there was a significant change in housing loan interest rates. I have 694.778 disbursed, 247.572 prepaid home loans between the specified period.

CHAPTER V

PREPAYMENT MODEL SELECTION

Borrowers typically have the option to prepay their mortgages at any time and most of the Banks in Turkey assume a constant prepayment rate for hedging purposes or valuation. The purpose of this thesis is to analyze the variables related to home loan prepayment and give an idea for the home loan market in Turkey. These variables are specified by loan level. Each loan consists of important information and preferred to work with the maximum information that can be used.

Two types of modeling are popular in prepayment studies. One of them is Cox Proportional Hazard Models and the other is Binary Models. The biggest difference between Cox and Logistic Regression is either included time or not. Hazard regression estimates the effects of variables and thus the loan survival duration. The pros and cons are shown in Figure 16 [27]. The survival model calculates the duration of the home loan until it dies (default or prepaid). Binary model outputs are the probability of prepayment on the loan or pool level. In this paper, the main aim is to focus on if the prepayment of the loan occurs or not. Furthermore, it is challenging to include time-varying loan-level covariates in survival models compared to logistic regression. For these reasons, logistic regression algorithm for the analysis of prepayment for home loans is selected.

	Model	Output	Positive Aspects	Negative Aspects
Proportional Hazard Models	Poisson regression model with multiple age-dependent constant baseline hazards (Schwartz, Torous, 1993)	Survival analysis (duration model)	Clear difference for prepayment rates between 'contract age' groups.	Multiple models and parameters are necessary to model age-dependency.
	A log-logistic baseline hazard function (He, Liu, 1998)		Intuitive interpretable modelling of time-dependency.	Multiple parameters are to be estimated through a complex probability density function and MLE procedure
	Constant baseline hazard (Cox,1972)		All dependency on explanatory variables is modelled in the exponent.	Outcome values not bounded between zero and one. Complex partial likelihood estimation.
Binary Response Models	Logistic Regression (Logit Model)	Discrete events. Prepayment probability by loan level. (or a prepayment probability on a pool or portfolio level)	All: Bounded outcome variable. Built-in Eviews and consequently Maximum Likelihood Estimation. No normality assumptions with respect to the residuals. Logit:Odds ratio makes output interpretable.	No disadvantages.
	Probit Regression			Both: More complex cdf than Logit model. No odds ratio or derivative calculations to interprete coefficients intuitively.
	Gumbel Regression			Gumbel: Meant to model extreme events.

Figure 16: Literature Summary, Advantages and Disadvantages of Modelling Types

Each model type has positive aspects, but the logistic regression type has no disadvantages for this study topic. Logistic regression is easy to use because of its assumptions such as each event is independent, there is no need to assume for regression residuals normality, it is enough that regression residuals are independent. The odds ratio helps to decide about each variable easily. Logistic regression models calculate the effects of independent variables on outcome (target variable). Risk factors calculated as a probability.

The better prediction with a Binary Model, logistic regression, there are two possible outcomes defined; the first one is “prepaid” and the second one is “payment of a regular installment.”

Prob (X=1) = p_i = probability of prepayment

Prob (X=0) = 1- p_i = probability of regular payment

With Bernoulli’s distribution, two outcomes occur in sample data. The logistic function on which logistic model is specified as follows:

$$f(Z) = \frac{1}{(1 + e^{-z})} \quad \text{(8)}$$

One unit change in the variable x, (Δx), is pointing to the ratio of the change in Probability of X (P(x)). The formula for logistic regression is as follows:

$$E(Y = \frac{1}{xi}) = \frac{1}{1 + \mathrm{e}^{-(\propto + \beta x)}} \quad \text{(9)}$$

Take as Z; $Zi = \alpha + \beta . Xi$; Then the logistic distribution function is as follows:

$$Pi = \frac{1}{1+e^{-zi}} \text{ and } 1 - Pi = \frac{1}{1+e^{zi}} \quad \text{(10)}$$

$Z_i = \{-\infty, +\infty\}$ and P_i between 0 and 1

The formula for odds ratio is as follows:

$$\frac{Pi}{1 - Pi} = \frac{1 + e^{zi}}{1 + e^{-zi}} = e^{zi} \quad \text{(11)}$$

Coefficients are $\beta = \beta_1, \beta_2, \beta_3, \ldots, \beta_n$

Explanatory variables are $x = x_1, x_2, x_3, \ldots, x_n$

The formula for prediction function is as follows:

$$\ln\left(\frac{Pi}{1 - Pi}\right) = \ln(e^{zi}) = Zi = \beta o + \beta 1.x1 + \beta 2.x2 + \cdots + \beta n.xn$$

$$\mathrm{P(x)} = \frac{1}{1 + e^{-(\beta 0 + \beta 1.x1 + \cdots + \beta n.xn)}} \quad \text{(12)}$$

In the literature, it is observed that logistic regression is preferred in many studies.

Tyrone T. Lin, Chia-Chi Lee, Ming-Sung Hu analyzed the default and the prepayment models using logistic regression dividing the factors as personal factors, contract terms and external factors in 2009 [28]. Their sample is minor, branch in Taiwan domestic bank, data includes only January 2002 and June 2007.

Jan-Yee Kung, Cheng-Chung Wu, Shih-Yun Hsu, Shiu-Shia Wu Lee and Chung-Wen Yang (2010) analyzed the home loans in Taiwan’s five commerce banks to explain the key factors that affect prepayments. As a result of the study, the logistic regression model can provide clear results to calculate prepayment probability. Leverage (available loan amounts), grace period (number of months), customer age, home loan duration, customer income (double income families are more likely) are significant variables [29].

Pedro Vasconcelos, (2010) used to explain prepayment risk with the multinominal Logit Model on a loan-level on the Dutch mortgage portfolio of the National Industrial Bank of China [24].

Xiaowei Li, Weida Kuang study loan-level data in China and focused on borrowers’ behavior and the segmentation analysis of the prepayment risk in 2012 [30]. Results from their logit model; housing valuation rate and market performance are key concepts for the Chinese customers.

CHAPTER VI

REGRESSION RESULTS

6.1 Summary Statistics

Detailed summary for each selected candidate variable in this section is provided. The quality of the data is examined by looking at the percentile values. Table 15 and Table 16 present selected variables and calculated percentile values. The missing rate of the variables in the sample is below 1,1%. No assumptions have been made regarding extreme values.

Table 15: Summary Statistics of Selected Variables (Loan and Borrower specific)

Variable	1st Pctl	5th Pctl	10th Pctl	25th Pctl	Median	75th Pctl	90th Pctl	95th Pctl	99th Pctl	Observation	Missing	% Miss
Loan Term	30	45	54	66	109	120	120	123	179	29.764.008	-	-
Current Principal Balance	3.072	10.909	18.213	34.492	59.398	97.446	159.685	216.000	392.484	29.743.322	20.686	0,07%
Current Principal Balance (LogX/CPI)	5,46	6,71	7,26	8,03	8,65	9,15	9,55	9,81	10,36	29.743.320	20.686	0,07%
Borrower Total Debt	3.242	12.751	20.774	38.088	64.886	106.477	177.173	241.107	448.257	29.738.983	25.025	0,08%
Borrower Total Debt (LogX/CPI)	5,52	6,87	7,41	8,14	8,74	9,23	9,64	9,90	10,49	29.738.983	25.025	0,08%
Total Risk	3.084	10.953	18.288	34.638	59.672	97.924	160.581	217.505	396.166	29.743.322	20.686	0,07%
Total Risk (LogX/CPI)	5,46	6,71	7,27	8,04	8,66	9,15	9,56	9,81	10,37	29.743.322	20.687	0,07%
Customer Age	24	27	29	34	41	50	57	61	67	29.762.028	1.980	0,01%
Interest Rate	0,64	0,66	0,71	0,78	0,89	0,99	1,09	1,13	1,32	29.764.008	-	-
First Loan Time	1	5	10	26	58	110	179	215	272	29.761.346	2.662	0,01%
Time with Bank	1	7	13	34	74	133	183	207	233	29.762.067	1.941	0,01%
Original Value	50.000	65.000	75.000	100.000	135.000	200.000	300.000	400.000	760.000	29.475.314	288.694	0,98%
Original LTV	0,2	0,31	0,39	0,53	0,68	0,75	0,8	0,8	0,9	29.475.308	288.700	0,98%
Interest spread	-1,50	-1,02	-0,75	-0,45	-0,23	-0,06	0,07	0,16	0,28	29.764.008	-	-
Interest ratio	0,35	0,44	0,52	0,66	0,79	0,94	1,07	1,18	1,37	29.764.008	-	-
Present Value Ratio	0,35	0,44	0,52	0,66	0,79	0,94	1,07	1,18	1,37	29.764.008	-	-
Current Penalty	31	110	191	529	1.117	1.900	3.135	4.242	7.720	29.743.322	20.686	0,07%
Current Penalty (LogX/CPI)	0,86	2,11	2,73	3,83	4,68	5,21	5,62	5,88	6,42	29.743.322	20.686	0,07%
Current LTV	0,13	0,21	0,25	0,35	0,46	0,59	0,70	0,74	0,81	29.472.969	291.039	0,99%
Current Value	61.057	83.582	100.020	135.987	193.748	288.000	428.627	559.016	1.037.052	29.472.969	291.039	0,99%
Customer Income	831	1.200	1.500	2.089	3.040	4.865	7.220	10.000	25.000	29.575.213	188.795	0,64%
Customer Income (LogX/CPI)	4,20	4,61	4,84	5,24	5,66	6,06	6,49	6,82	7,76	29.575.213	188.796	0,64%
Current Penalty to Current Value Ratio	0,000	0,000	0,001	0,003	0,006	0,010	0,013	0,014	0,016	29.439.427	324.581	1,10%
Prepayment Option Value	-0,602	-0,537	-0,373	-0,325	-0,117	0,012	0,077	0,209	0,316	29.743.320	20.688	0,07%
Current Penalty to Customer Income Ratio	0,007	0,029	0,053	0,142	0,367	0,571	0,766	0,882	1,135	29.555.337	208.671	0,71%
Loan Age	0	2	4	11	26	48	71	84	104	29.764.008	-	-
Customer Education Status	1	2	2	2	3	4	5	5	6	29.762.089	1.919	0,01%
Salary Customer Flag	1	1	1	1	2	2	2	2	2	29.762.089	1.919	0,01%

Table 16: Summary Statistics of Selected Variables (Macro-economic Variables)

Variable	1st Pctl	5th Pctl	10th Pctl	25th Pctl	Median	75th Pctl	90th Pctl	95th Pctl	99th Pctl	Observation	Missing	% Miss
RPPI - Level	48,8	53,6	58,4	75,4	101,4	120,9	165,5	192,7	247,4	29.764.008	-	-
RPPI - (m-m, %)	-0,008	0,002	0,004	0,007	0,011	0,016	2,350	3,910	9,039	29.761.376	2.632	0,01%
Reference Interest Rate	8,77	9,47	10,73	11,42	13,06	16,86	18,19	22,98	28,63	29.764.008	-	-
Sector Housing Loan Amount; Level	61.520	74.992	91.303	128.000	180.986	202.694	277.346	280.037	298.896	29.764.008	-	-
Sector Housing Loan Amount (LogX/CPI)	8,840	9,022	9,134	9,435	9,684	9,827	9,921	10,023	10,069	29.764.008	-	-
Sector Housing Loan Amount; Level, yearly change (%)	-9,81	-5,01	0,34	9,56	14,90	21,56	36,55	43,46	51,57	29.764.008	-	-
Share of Housing Loans in total consumer loans (%)	38,34	38,60	39,83	42,17	44,46	47,28	48,29	48,77	49,12	29.764.008	-	-
CPI	6,13	6,58	7,20	8,10	10,61	15,01	19,58	20,35	36,08	29.764.008	-	-
USD/TRY Level	1,56	1,79	1,83	2,46	3,68	6,06	8,02	8,61	13,55	29.764.008	-	-
USD/TRY (LogX/CPI)	-1,78	-1,52	-1,46	-1,27	-1,04	-0,79	-0,61	-0,48	-0,41	29.764.008	-	-
USD/TRY yearly change (%)	-7,25	-1,28	1,87	10,28	18,55	27,52	37,90	41,48	75,22	29.764.008	-	-
YCS	-6,83	-0,53	0,14	0,78	1,05	1,26	1,69	2,82	17,20	29.764.008	-	-
Total Industry-Level	73,39	78,88	84,84	95,08	108,69	119,62	133,07	142,17	165,56	29.764.008	-	-
Capacity Utilization Rate of Manufacturing Industry	62,60	73,30	74,10	75,30	76,70	77,90	78,50	78,80	79,70	29.764.008	-	-
Unemployment Rate	7,60	8,40	9,00	10,00	10,90	12,90	13,50	14,10	15,00	29.764.008	-	-
TRY2Y-end	5,61	7,43	8,35	9,08	11,11	15,52	18,90	21,71	23,93	29.764.008	-	-
TRY10Y-end	6,55	7,96	8,56	9,54	10,85	13,85	17,73	18,95	23,05	29.764.008	-	-
Divorce Rate, annual	1,59	1,59	1,60	1,64	1,69	1,76	2,07	2,07	2,07	29.764.008	-	-
Sector Housing Loan Amount; monthly change (%)	-0,14	-0,14	-0,14	-0,01	0,00	0,08	0,26	0,26	0,26	29.764.008	-	-
USD/TRY monthly change (%)	-0,08	-0,04	-0,02	-0,01	0,01	0,04	0,07	0,08	0,29	29.764.008	-	-
Divorce Rate,change	-0,01	-0,01	-0,01	0,00	0,01	0,02	0,02	0,03	0,12	29.764.008	-	-

Table 17: Summary Statistics

Parameters	Non-prepaid		Prepaid		Total	
	Mean	SD	Mean	SD	Mean	SD
Number of Loans	447.206	-	247.572	-	694.778	-
Current LTV	0,57	80,19	1,50	310,89	0,89	193,82
Interest Rate	0,94	0,19	0,99	0,22	0,94	0,19
Loan Term	93,73	35,32	93,85	32,82	93,73	35,32
Loan Age	44,51	32,43	36,90	27,31	41,83	30,94
Customer Age	39,25	10,17	42,13	10,88	39,25	10,17
Borrower Total Debt	124.758	175.157	32.697	98.470	94.377	160.091
Interest spread	-0,45	0,26	-0,15	0,35	-0,34	0,33

Table 17 displays selected descriptive statistics. Average term loan is 93 months and there is no difference between prepaid and non-prepaid loans. Average non-prepaid loan age is 44,5 months higher than the prepaid loans with the average loan age 36,9 months. The average interest rate for prepaid loans is 0,99% and it is higher than the rate of 0,94% non-prepaid loans. Average non-prepaid customer age is 3 years smaller than the prepaid loans with the average customer age 42 years. This is the opposite of what is expected. Non-prepaid customers' total debt is 3,8 times that of prepaid customers' total debt.

6.2 Regression Model Results

19 macroeconomic and 23 loan and customer specific variables are selected to examine the effect of factors on home loan prepayment between the period 2010 and 2021. In Table 18 the parameters are shown with standard errors, standardized estimates. Analysis of Maximum Likelihood Estimates show that the variables that may be significant such as Total Industry-Level, Current LTV, Customer Income are not in the table.

Table 18: Analysis of Maximum Likelihood Estimates

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq	Standardized Estimate
Intercept	1	-6,248	0,5761	117,63	<,0001	
Interest spread	1	-15,522	0,0974	25.418,98	<,0001	-3,038
Reference Interest Rate	1	-0,933	0,0075	15.648,84	<,0001	-2,1226
Total Risk; log (x/CPI)	1	3,747	0,5500	46,42	<,0001	1,9907
Borrower Total Debt; log (x/CPI)	1	-2,582	0,0054	231.693,28	<,0001	-1,4604
Current Principal Balance; log (x/CPI)	1	-1,712	0,5514	9,64	0,0019	-0,909
Present Value Ratio	1	6,832	0,1033	4.377,29	<,0001	0,8191
Current Penalty to Customer Income Ratio	1	-0,596	0,0228	680,61	<,0001	-0,6217
Interest Rate	1	7,443	0,0792	8.834,44	<,0001	0,6068
Interest ratio	1	4,987	0,1210	1.699,14	<,0001	0,6002
TRY10Y-end	1	-0,297	0,0062	2.304,98	<,0001	-0,5854
USD/TRY yearly change (%)	1	-0,039	0,0006	4.006,78	<,0001	-0,3265
TRY2Y-end	1	0,126	0,0052	585,26	<,0001	0,3102
Divorce Rate, annual	1	2,850	0,1507	357,62	<,0001	0,2457
Loan Term	1	0,013	0,0002	6.654,03	<,0001	0,2214
CPI	1	0,076	0,0047	258,76	<,0001	0,2183
Customer Income; log (x/CPI)	1	0,494	0,0038	16.961,32	<,0001	0,1934
USD/TRY Level, log (x/CPI)	1	1,080	0,0476	515,72	<,0001	0,191
Unemployment Rate	1	-0,192	0,0040	2.332,89	<,0001	-0,1869
USD/TRY monthly change (%)	1	6,586	0,1372	2.303,28	<,0001	0,1818
Customer Age	1	-0,030	0,0004	7.080,04	<,0001	-0,1727
Salary Customer Flag	1	-0,618	0,0070	7.804,23	<,0001	-0,1621
Sector Housing Loan Amount; monthly change (%)	1	-2,502	0,1770	199,62	<,0001	-0,1533
Sector Housing Loan Amount; log (x/CPI)	1	-0,899	0,0526	291,64	<,0001	-0,1478
Current Penalty ; log (x/CPI)	1	-0,215	0,0169	160,44	<,0001	-0,1383
RPPI - Level	1	0,005	0,0007	53,61	<,0001	0,1163
Original LTV	1	0,000	0,0000	317,68	<,0001	-0,1014
Prepayment Option Value	1	-0,756	0,0252	896,86	<,0001	-0,0889
Divorce Rate, annual change	1	-7,674	0,2480	957,51	<,0001	-0,0795
Capacity Utilization Rate of Manufacturing Industry	1	0,052	0,0018	849,63	<,0001	0,0786
Loan Age	1	-0,005	0,0002	435,62	<,0001	-0,0701
First Loan Time	1	0,002	0,0001	487,70	<,0001	0,0705
Time with Bank	1	0,001	0,0001	146,51	<,0001	0,0402
Sector Housing Loan Amount; Level, yearly change (%)	1	0,006	0,0005	148,20	<,0001	0,0408
Current Penalty to Current Value Ratio	1	0,009	0,0013	51,35	<,0001	0,0306
Share of Housing Loans in total consumer loans (%)	1	0,015	0,0040	14,51	0,0001	0,0257
RPPI - (m-m, %)	1	0,024	0,0074	10,65	0,0011	0,021
Customer Education Status	1	0,029	0,0028	104,61	<,0001	0,0195

Some metrics such as Gini, AUC (Area Under the Curve) from the ROC (Receiver Operating Characteristics) Curve and Somers'D are used as model's discriminatory power in the literature. The Gini and Somers'D coefficients are ratios that show how close our model is to being a "perfect model" and how far it is to be a "random model". The probability prepayment function that perfectly separates prepayment and (continue) pay on maturity borrowers, has a Gini index/ Somers'D equal to 1. The regression results, shown in Table 19 are satisfactory with 0,843 Somers'D index. ROC curve is a performance measurement. The higher the AUC from the ROC Curve, the better the model discriminates between classes. 0,921 AUC means that the model distinguishes prepaid and non-prepaid home loans. ROC curve is shown in Figure 17.

Table 19: Association Table

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	88,7	Somers' D	0,843
Percent Discordant	4,4	Gamma	0,906
Percent Tied	7,0	Tau-a	0,012
Pairs	6,13E+12	ROC Curve (Area)	0,921

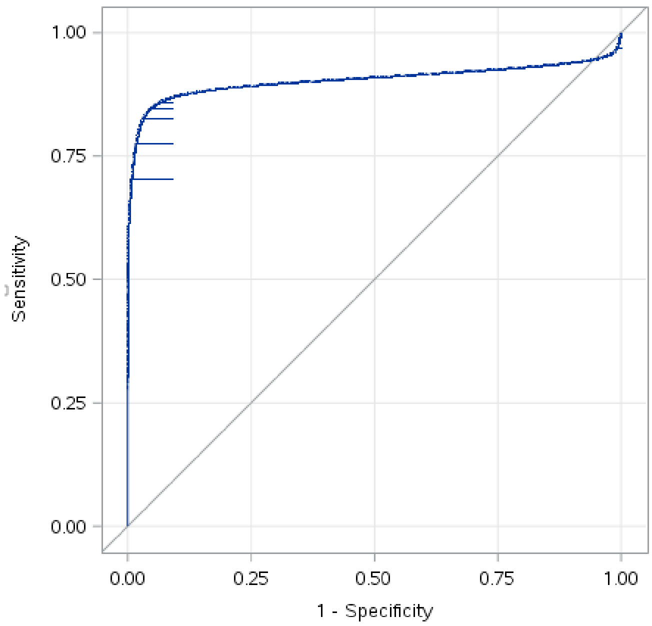

Figure 17: ROC Curve for Model

According to empirical findings, the most important factor affecting prepayment is the interest spread (negative). The second factor is the reference interest rate (market rate) (negative). After that respectively total risk of the home loan (positive), borrower total debt (negative), current principal balance of the home loan (negative) and PVR ratio (positive). Table 20 shows the explanatory effect of each variable. Approximately 78% of the prepayment probability of the model is explained by the first 10 variables.

Table 20: Univariate Gini (Order of the Independent Variables)

Parameter	**Univarite Gini**	**SİGN**
Interest spread	18,54%	-
Reference Interest Rate	12,95%	-
Total Risk; log (x/CPI)	12,15%	+
Borrower Total Debt; log (x/CPI)	8,91%	-
Current Principal Balance; log (x/CPI)	5,55%	-
Present Value Ratio	5,00%	+
Current Penalty to Customer Income Ratio	3,79%	-
Interest Rate	3,70%	+
Interest Ratio	3,66%	+
TRY10Y-end	3,57%	-
USD/TRY yearly change (%)	1,99%	-
TRY2Y-end	1,89%	+
Divorce Rate, annual	1,50%	+
Loan Term	1,35%	+
CPI	1,33%	+
Customer Income; log (x/CPI)	1,18%	+
USD/TRY Level, log (x/CPI)	1,17%	+
Unemployment Rate	1,14%	-
USD/TRY monthly change (%)	1,11%	+
Customer Age	1,05%	-
Salary Customer Flag	0,99%	-
Sector Housing Loan Amount; monthly change (%)	0,94%	-
Sector Housing Loan Amount; log (x/CPI)	0,90%	-
Current Penalty ; log (x/CPI)	0,84%	-
RPPI - Level	0,71%	+
Original LTV	0,62%	-
Prepayment Option Value	0,54%	-
Divorce Rate, annual change	0,49%	-
Capacity Utilization Rate of Manufacturing Industry	0,48%	+
First Loan Time	0,43%	+
Loan Age	0,43%	-
Sector Housing Loan Amount; Level, yearly change (%)	0,25%	+
Time with Bank	0,25%	+
Current Penalty to Current Value Ratio	0,19%	+
Share of Housing Loans in total consumer loans (%)	0,16%	+
RPPI - (m-m, %)	0,13%	+
Customer Education Status	0,12%	+

The interest spread is the most important variable affecting prepayment. Interest spread is defined as difference between contract rate and market rate. As the market interest rate decreases, the spread increases. As the spread increases, the probability of closing will increase as expected. And as most expected, the prepayment rate increases as reference (market) interest rate falls. Since interest rate changes can have linear or exponential effects, derivatives of the interest rate variable are taken. Interest ratio is defined contract rate divided by current market rate. As the contract rate becomes higher, the ratio increase and the prepayment rate also increase as expected. Interest rate of the loan has positive effect. The higher the contract interest rate, the lower the market rate makes it rational to prepay the loan. A lower coupon means a lower prepayment speed[31].

The other most important variable is total risk of the loan. The sign is positive and it can be interpreted that the higher the total risk of the loan, the more the factors affect it. Total risk of the loan is more vulnerable to the macro-economic and other factor changes. The total risk of the loan includes the amount subject to early payment penalty. The higher the risk, the higher the penalty occurs. This means; the higher the effect will be calculated for rational economic prepayment. Total risk includes current principal balance and accrued interest rates and insurance costs.

The borrower's total debt is the other important variable. The higher the customer indebtedness ratio, the lower the probability of prepaying. The customer with more total debt can close the mortgage loan with less probability.

The current principal balance is another explanatory variable of the model. The higher current principal balance means the less prepayment probability. As the principal balance rises, prepaying the loan may become more difficult to borrower. Total risk and current principal balance are entered the model as explanatory variable with reverse sign.

The accrued interest, which is included in the total risk variable, is the key component that causes opposite sign.

Derived variable, Present Value Ratio, is another explanatory variable that has a significant positive impact on the model. Ratio includes the reference market home loan interest rates and the contract interest rate and also includes remaining term of the loan.

Current penalty to customer income ratio has negative significant influence on prepayment probability. If the prepayment penalty is much higher than the income declared by the customer, the possibility of prepayment of the customer decreases again. The customer is expected to have a higher income and thus more easily to compensate for the prepayment penalty.

It is seen that the higher the USD/TRY level the higher the probability of prepayment. Also, the USD/TRY variable has a high impact on the CPI (exchange rate pass-through on inflation). Similarly, the increase on the two variables have positive effect (increase the prepayment probability). The higher the USD/TRY and CPI level, the higher the probability of prepay. High exchange rate and high CPI indicates the depreciation of TL. Since the depreciation of TL may cause a relative decrease in the loan debt balance, it can be interpreted that some customers may close their loans even if it is not rational.

The effects of house original and current values on prepayment are almost zero. As the value of the house increases, borrowers do not sell their houses and seek to buy a better house, as in other housing markets. The time of being a bank customer or customer education, which can be an indicator of financial literacy, is taken into account, the effect on prepayment is close to zero again. Contrary to studies conducted outside of Turkey, the age of the customer and loan age are also irrelevant to the prepayment probability in our sample.

There is always error in any empirical model. A variable that explains the prepayment behavior may have been left out. This risk will always be present in behavioral models, as people's behavior will not always be optimal. Logit model is used when two possible outcomes occurs. However, there are three distinguished outcomes for home loans; prepayment, default and continue. In conclusion, since the default rate is very low, it would not be wrong to set it up as 2 outcomes.

CHAPTER VII

CONCLUSION

Prepayment makes it difficult to manage the bank's balance sheet by disrupting the contractual cash flows, especially due to the meltdown in the loan portfolio. Banks try to manage their risks by booking derivatives and such transactions.

The thesis aims to explain the variables that might affect the prepayment event of fixed rate home loans in Turkey. With Area-under-the-Curve (AUC) statistic of 0.921 and Gini coefficient of 0.843, regression fit the model with selected macro-economic, borrower specific and loan-related variables. Empirical results indicate that the interest rate change in the market and the total debt of the customer and solvency are the factors that affect the probability of prepayment the most. In Turkey, it is observed that housing loans are prepaid with the decrease in market interest rates.

Compared with the papers outside Turkey, it is observed that customers in Turkey have different behaviors except for the interest variable. The most important difference is that the LTV ratio and the increase in the price of the house have almost no effect. Similarly, customer age and income and geographic location have very minor effects. Although this thesis includes the regional housing price index (RPPI), the effect is negligible. Since the current value of the house has no effect on prepayment, it is seen that the house price-earnings ratio is not taken into account. Between 2010 and 2021, borrowers in Turkey do not sell their houses for profit.

The model can be further extended and made more robust by adding internal or external information regarding the credit capacity of borrower. For instance, credit scores and Check and Risk reports supplied by the Turkish Credit Registration Bureau may provide rich information regarding along this dimension.

BIBLIOGRAPHY

[1] "TUIK." https://www.tuik.gov.tr/ (accessed Jan 08, 2019 and accessed May 5, 2022)

[2] "EVDS." https://evds2.tcmb.gov.tr/index.php?/evds/serieMarket (accessed May 5, 2022)

[3] "On Birinci Kalkınma Planı (2019-2023): Konut Politikaları Özel İhtisas Komisyonu Raporu," Ankara, 2018. [Online]. Available: https://www.sbb.gov.tr/wp-content/uploads/2020/04/KonutPolitikalariOzelIhtisasKomisyonuRaporu.pdf (accessed May 5, 2022)

[4] "Mortgage Prepayment." http://www.rbcroyalbank.com/mortgages/pdf/MortgagePrePaymemt_E.pdf (accessed May 5, 2022)

[5] *Tapu Kanunu ve Kadastro Kanununda Değişiklik Yapılmasına İlişkin Kanun*. [Online]. Available: https://www.resmigazete.gov.tr/eskiler/2012/05/20120518-1..htm (accessed May 5, 2022)

[6] "İkamet İzinleri." https://www.goc.gov.tr/ikamet-izinleri (accessed May 5, 2022)

[7] "Dokuzuncu Kalkınma Planı (2007-2013)," 2006. [Online]. Available: https://www.sbb.gov.tr/wp-content/uploads/2021/12/Dokuzuncu_Kalkinma_Plani-2007-2013.pdf (accessed May 5, 2022)

[8] "Onuncu Kalkınma Planı (2014-2018)," Ankara, 2013. [Online]. Available: https://www.sbb.gov.tr/wp-content/uploads/2021/12/Onuncu_Kalkinma_Plani-2014-2018.pdf (accessed May 5, 2022)

[9] "On Birinci Kalkınma Planı (2019-2023)," Ankara, 2019. [Online]. Available:

https://www.sbb.gov.tr/wp-content/uploads/2021/12/On_Birinci_Kalkinma_Plani-2019-2023.pdf (accessed May 5, 2022)

[10] *Türk Vatandaşlığı Kanununun Uygulanmasına İlişkin Yönetmelikte Değişiklik Yapılmasına Dair Yönetmelik*. [Online]. Available: https://www.resmigazete.gov.tr/eskiler/2018/09/20180919-15.pdf (accessed May 5, 2022)

[11] *Hazineye ait Taşınmaz Malların Değerlendirilmesi ve Katma Değer Vergisi Kanununda Değişiklik Yapılması Hakkında Kanun ile Bazı Kanunlarda ve Kanun Hükmünde Kararnamelerde Değişiklik Yapılmasına Dair Kanun*. [Online]. Available: https://www.resmigazete.gov.tr/eskiler/2022/04/20220415.pdf (accessed May 5, 2022)

[12] *Tüketicinin Korunması Hakkında Kanun*. 2013. [Online]. Available: https://www.resmigazete.gov.tr/eskiler/2013/11/20131128-1.htm (accessed May 5, 2022)

[13] S. Amar, "Modeling of Mortgage Loan Prepayment Risk with Machine Learning," Technical University of Munich.

[14] A. Goodarzi, R. Kohavi, R. Harmon, and A. Senkut, "Loan Prepayment Modeling," *Am. Assoc. Artif. In- telligence*, 1998.

[15] J. C. Hu, *Appendix A.: Analysis of Prepayment and Prepayment Rate*. Wiley, 2011.

[16] "BDDK." https://www.bddk.org.tr/BultenHaftalik/ (accessed May 5, 2022)

[17] "Principles for the Management of Interest Rate Risk," 1997. [Online]. Available: https://www.bis.org/publ/bcbs28a.pdf (accessed May 5, 2022)

[18] "Interest Rate Risk." https://www.fdic.gov/resources/bankers/capital-

markets/interest-rate-risk/IRR-2.PDF (accessed May 5, 2022)

[19] "Interest Rate Risk in the Banking Book," 2016. [Online]. Available: https://www.bis.org/bcbs/publ/d368.pdf (accessed May 5, 2022)

[20] A. Castagna and F. Fede, *Measuring and Managing Liquidity Risk*. Wiley, 2013.

[21] P. Chinloy, "The Option Structure of a Mortgage Contract," *J. Hous. Res.*, vol. 2, no. 1, pp. 21–38, 1991.

[22] R. Doff, "Behavioural Risk Management in the Financial Markets," in *Behavioural Risk Management*, Risk.net, 2020.

[23] M. Formenti and U. Crespi, *A Guide to Behavioural Modelling*. 2019.

[24] P. Vasconcelos, "Modelling Prepayment Risk: Multinominal Logit Model Approach For Assesing Conditional Prepayment Rate," University of Twente, 2010.

[25] S. Van Nieuwerburgh, M. Chernov, B. R. Dunn, and F. A. Longstaff, Eds., "Macroeconomic-Driven Prepayment Risk and the Valuation of Mortgage-Backed Securities," *Rev. Financ. Stud.*, vol. 31, no. 3, pp. 1132–1183, 2018.

[26] J. Meis, "Modelling Prepayment Risk in Residential Mortgages," Erasmus University Rotterdam, 2015.

[27] C. T. J. Hamstra, "Floating-rate Mortgages: Why Do They Prepay?," University of Twente, 2010. [Online]. Available: http://essay.utwente.nl/63017/1/Scriptie_Stijn_Hamstra_Floating_rate_mortgage_prepayments_ING_Bank.pdf (accessed May 5, 2022)

[28] T. T. Lin, C.-C. Lee, and M.-S. Hu, "The Determinants of Mortgage Defaults and Prepayments," *J. Inf. Optim. Sci.*, vol. 31, no. 5, pp. 993–1009, 2010.

[29] J.-Y. Kung, C.-C. Wu, S.-Y. Hsu, S.-S. W. Lee, and C.-W. Yang, "Application of Logistic Regression Analysis of Home Mortgage Loan Prepayment and Default

Risk," *ICIC Express Lett.*, vol. 4, no. 2, pp. 325–331, 2010, [Online]. Available: https://ir.csu.edu.tw/dspace/bitstream/987654321/3004/2/Application+of+Logistic+Regression+Analysis+of+Home+Mortgage+Loan+Prepayment+and+Default+Risk.pdf (accessed May 5, 2022)

[30] X. Li and W. Kuang, "Prepayment Behavior of China's Mortgage Borrowers: Evidence from the Loan-Level Data," *Int. J. Hous. Mark. Anal.*, vol. 5, no. 1, pp. 71–88, 2012.

[31] J. P. A. M. Jacobs, R. H. Koning, and E. Sterken, "Modelling Prepayment Risk," 2005. [Online]. Available: https://www.researchgate.net/profile/Jan-Jacobs-5/publication/266568317_MODELLING_PREPAYMENT_RISK/links/54e47ed80cf282dbed6fe41c/MODELLING-PREPAYMENT-RISK.pdf (accessed May 5, 2022)

VITA

After graduating from Economics at Ankara University in 2005, she started her career as an assistant specialist in a dealing-room of the State Bank in Treasury Department. After working at foreign currency money market desk for 2.5 year, she has worked for 6 years as an auditor in the Board of Audit. Throughout her years at audit, she has learned operational procedures of Bank, risk management in branches, control functions, and governance processes in Bank. Then she transferred to Treasury Department again. She widen her viewpoint in asset and liability management in Treasury Department for 2 years.

After her experience in bank fund management and audit, she was appointed to the risk management department as the manager of the validation team. After working as a validation manager for 2 years, she has been working as a market risk management manager for 3 years. Throughout her years at market risk, she learned how to calculate VaR, actively played a role analyzing liquidity, interest rate risk and foreign exchange risk. She closely monitors balance sheet risks and takes active role in stress tests, ICAAP, recovery plan studies. She has 15 years of experience in different units and positions.