

Department of Management and Production Engineering

Master's Degree in "Engineering and Management"

Finance Field

Master's Degree Thesis

Italians' choice between fixed and variable interest rate in the mortgage market: an empirical investigation.

Supervisor

Candidate

Prof. Riccardo Calcagno

Giulia Rostagno

Academic Year 2024/2025

Abstract

This thesis empirically analyses the determinants of the choice between fixed-rate mortgages and adjustable-rate mortgages in the Italian market, focusing on the period 2001-2010. The analysis is based on data collected from surveys conducted by the Bank of Italy, which enabled an examination of individual and family characteristics and their interaction with the market conditions.

The empirical model, developed using logistic regression, reveals that a higher initial mortgage amount and an extended mortgage duration are associated with a slight decrease in the probability of opting for a variable rate. The historical context plays a decisive role: while the period 2001-2005 is characterized by a predominance of fixed-rate choices, the period following the global financial crisis (2008-2010) favours the adoption of a variable rate, highlighting that expectations regarding future rate trends significantly influence borrowers' decisions. An unexpected result concerns risk attitude, whose influence appears counterintuitive: lower risk aversion is associated with a propensity toward a fixed rate.

Despite some limitations related to the specificity of the sample and the nature of the data employed, the study provides relevant insights for further research, contributing to a deeper understanding of the complex decision-making process in the mortgage market.

Index

Introduction	1
Chapter 1 - Literature Review	3
1.1 Theoretical background	3
1.2 Empirical studies	5
Study 1: "Household Risk Management and Optimal Mortgage Choice" by J Campbell and João F. Cocco, 2003	'ohn Y. 5
Study 2: "Why do Italian households prefer adjustable-rate mortgages?" by Zocchi, 2013	Paola 6
Study 3: "Mortgage Choice in the Euro Area: Macroeconomic Determinants the Effect of Monetary Policy on Debt Burdens" by Michael Ehrmann and M Ziegelmeyer, 2017	and 'ichael 7
1.3 Gaps and limitations	9
1.4 Relation with my analysis	10
Chapter 2 - Methodology & Data	13
2.1 Data source	13
2.2 Data selection	13
2.3 Data transformation	15
2.4 Description of variables	16
Variables related to the reference person	16
Variables related to the household	18
2.5 Descriptive statistics	21
Descriptive statistics related to the reference person	21
Descriptive statistics related to the household - Quantitative variables	23
Descriptive statistics related to the household - Qualitative variables	25
Chapter 3 - Results analysis	29
3.1 Correlation matrix	29
3.2 Logistic regression	31
Multicollinearity	31
Theoretical background	31
Model parameters	33
Standardized parameters	42
Model performance	44
Chapter 4 - Conclusions	47
4.1 Main results	47

4.2 Limitations	48
4.3 Further studies	49
4.4 Final considerations	50
Appendix A	51
A.1 Descriptive statistics at person level - Quantitative variables	51
A.2 Descriptive statistics at person level - Qualitative variables	52
A.3 Descriptive statistics at family level - Quantitative variables	53
A.4 Descriptive statistics at family level - Qualitative variables	54
Appendix B	55
Correlation matrix	55
Bibliography	56

Introduction

The decision to take out a mortgage represents an important milestone in the lives of many families. This choice can arise from different initial conditions and varied intentions, resulting in each individual and household having distinct needs and concerns. It is precisely this multiplicity of conditions that gives rise to the interest in studying the optimal choices to adopt based on the specific characteristics of each case. One of the most relevant decisions is the choice between an adjustable-rate and a fixed-rate mortgage. This option is influenced not only by the intrinsic characteristics of the family, but also by market conditions, both current and forecasted. On one hand, a fixed rate offers security and stability, while on the other hand, the flexibility of a variable rate can become advantageous or disadvantageous depending on the circumstances.

The literature shows that the choice between a fixed-rate mortgage (FRM) and an adjustable-rate mortgage (ARM) depends primarily on risk aversion, income stability, and the market rates. Households characterized by uncertain incomes or those seeking to avoid potential increases in instalments tend to favour FRMs, as they offer greater predictability of payments. In contrast, those with more stable or growing income generally exhibit a greater propensity towards ARMs, attracted by the lower initial rate. The size and duration of the mortgage are also influential: high amounts and extended terms, while allowing for a spreading out of payments over time, can make ARMs more attractive due to their lower initial instalments, although this exposes borrowers to potential future rate increases. Moreover, in periods of robust economic growth and low unemployment volatility, an ARM appears more sustainable, as families place their confidence in the continuity of their income. Finally, financial education plays a crucial role: a limited understanding of market mechanisms can lead to suboptimal choices, resulting in an underestimation of the future risks associated with an ARM.

The objective of this thesis is to concretely analyse the choice, between a variable and a fixed rate, made by Italian families between 2001 and 2010. During this time interval, there was, on one hand, relative rate stability with a pre-crisis increase (2001-2008) and, on the other hand, a drastic decline due to the financial crisis (2008-2010). These macro factors have certainly affected the availability of credit over the years.

The data used were entirely derived from surveys conducted by the Bank of Italy. Among the numerous variables collected, this thesis will focus on those related to the reference person (Age, Marital status, Level of education, and Occupational status) as well as those related to the household: Purpose, Year of signing, Initial amount, Duration, Current value all properties, Total real assets, Outstanding balance of household's liabilities, Expenses compared to income (last 12 months), Future income expectations and Risk attitude.

In summary, the analysis revealed that a higher initial amount and a longer duration are associated with a slight reduction in the probability of opting for a variable rate, with relatively significant impacts. A higher value of real assets, an indicator of economic stability, is instead correlated with a greater propensity towards a variable rate. Moreover, mortgages taken out during the period 2001-2005 tend to be FRMs, whereas those concluded between 2008 and 2010 show a preference for a variable rate, highlighting the influence of historical conditions. Finally, risk attitude does not influence the choice of interest rate as expected; rather, lower risk aversion appears to be associated with a tendency towards a fixed rate.

The following is a description of the structure of the thesis.

The first chapter is dedicated to the literature review and is divided into a brief theoretical introduction to the topic, a section that summarizes the subject and the main results obtained from the three articles, a part that highlights the main gaps and limitations common to these texts, and a final section that illustrates the relationships between these articles and this thesis.

The second chapter is related to methodology and data. It describes the data used and the methodology adopted, including the reasoning and any simplifications made, before presenting the descriptive statistics.

The third chapter presents the results obtained from the analysis using logistic regression, interpreting them and comparing them with what is reported in the literature and with the initial expectations; in particular, it illustrates the correlations between the model's variables, analyses the parameters (both standardized and non-standardized), and assesses the model's performance.

The final chapter represents the conclusion of the thesis, in which not only the key points of the analysis are highlighted, but also the limitations encountered and the possible improvements are discussed, while also providing suggestions for future studies.

Chapter 1 - Literature Review

1.1 Theoretical background

The economic and financial literature has devoted considerable attention to the choice between FRMs and ARMs, as this decision represents a crucial component of the household's financial risk management. In general terms, theoretical studies frame the issue as a choice that involves balancing the predictability of payments with the potential variability of debt costs over time.

On one hand, a FRM guarantees cash flow stability for the household, protecting it from uncertainty regarding future interest rates. This characteristic helps avoid unexpected shocks to monthly payments and facilitates long-term financial planning. On the other hand, an ARM more closely aligns payments with current short-term market interest rates. This structure allows the borrower to benefit from any interest rate decreases but exposes the family to higher risks if rates increase.

FRMs can potentially become costly in real terms when inflation decreases, since payments, fixed in nominal terms, do not decrease in real value. When inflation rises, however, FRMs become cheaper in real terms. ARMs do not present this issue, as payments adjust to market interest rates; it can be therefore said that they offer a more stable real capital value.

The main structural difference between the two lies in the interest rates, which are based on different reference index. For FRMs, the interest rate typically used is the IRS (Interest Rate Swap) for medium-to-long-term periods, matching the duration of the mortgage. It is referred to as "Eurirs" because the swap interest rate is calculated by taking into account the banks of the European Community that refer to the FBE (European Banking Federation). ARMs, instead, use the Euribor (Euro Interbank Offered Rate), which is the short-term rate at which the most reliable banks (FBE-member banks) offer interbank term deposits. Alternatively, ARMs can be based on the ECB (European Central Bank) rate. The interest rate applied to both types of mortgages is calculated using the following formula: Mortgage interest rate = Reference index + Spread.

The spread comprises various components: the borrower's default risk, the mortgage term, additional services provided by the bank, and the bank's profit strategy. For FRMs it also includes the cost of hedging instruments against upward interest rate movements,

compensating for the risk that the borrower might decide to terminate the contract in order to refinance under more favourable conditions.

In many markets, FRMs include a refinancing option that allows the borrower to reduce the debt cost if market interest rates fall below those available at the time of loan origination. This option, however, is not free, as FRMs rates typically incorporate a premium associated with the added flexibility. This premium is reflected in an initial interest rate generally higher than that of ARMs, making FRMs more expensive if inflation remains stable or decreases. In essence, the FRM involves a trade-off: it protects against future rate uncertainties but comes with a higher initial cost and greater sensitivity to the inflation environment.

However, if short-term rates are very high (for example, due to restrictive monetary policies), the variable rates of an ARM could exceed those of a long-term FRM. This scenario can occur under specific economic conditions, such as when investors expect short-term rates to decline in the future even though they are currently high. For this reason, the choice between fixed and variable rates is influenced by the differential between long-term and short-term rates (the yield spread), which determines the difference between FRMs and ARMs. When this yield spread is unusually high (long-term yield much higher than short-term ones), more homeowners should take out ARMs; when it is unusually low, more homeowners should take out FRMs.

A third, less common theoretical category, is inflation-indexed mortgages, which link payments to movements in the consumer price index. This approach aims to protect the lender from the loss of the money's purchasing power while providing the borrower with a more stable payment profile in real terms. Although theoretically appealing, this solution faces practical limitations, as it requires mature financial markets and tools capable of accurately measuring and managing inflation expectations.

Ultimately, the theoretical framework outlines an intertemporal optimization problem in which households, considering their own characteristics (income profiles, risk attitude, potential default costs, and the time horizon for holding the property) and the macroeconomic context (the level and volatility of inflation, the structure of short- and long-term interest rates, and credit market conditions), choose the type of mortgage that maximizes their expected welfare.

1.2 Empirical studies

Below there are three cases analysed in detail. All three concern the optimal choice of mortgage type, but each has different conditions and perspectives.

Study 1: "Household Risk Management and Optimal Mortgage Choice" by John Y. Campbell and João F. Cocco, 2003.

In their study, Campbell and Cocco propose a model for the optimal choice between FRMs and ARMs in a context characterized by inflation uncertainty and liquidity constraints. The authors structure their analysis using a life-cycle model, which is a dynamic framework that takes into account how the relevant variables evolve over the entire duration of the mortgage contract. This approach is more comprehensive than a static model, which merely offers a snapshot of the borrower's situation at a single point in time, for example, when the contract is signed.

Specifically, the analysis examines various key factors, including the household's risk profile and risk aversion, the variability of expected incomes, the size of the debt, and the flexibility of financial markets. This makes it possible to understand how households balance stability, credit accessibility, refinancing costs, and default risk over an extended time horizon.

The findings show that households with unstable or uncertain incomes, high risk aversion, and significant existing debts tend to prefer FRMs. This preference arises from the predictability of payments over time: the certainty of a constant payment, regardless of market interest rate fluctuations, reduces the probability of default, even in scenarios where rates rise suddenly. In contrast, households that are more tolerant of cost variability, such as those with stable or growing future income prospects or less-constraining debt levels, may find ARMs more appealing. This comes from the fact that they generally offer lower initial rates and, potentially, lower overall costs if short-term rates decrease over time.

Another crucial aspect highlighted by the research is the importance of the spread between long-term rates (influencing FRMs) and short-term rates (affecting ARMs). If the spread between long- and short-term bond yields is large, an ARM appear relatively more attractive. On the contrary, when this differential reduces or even inverts, indicating a flattening of the yield curve and a reduced gap between long- and short-term rates, FRMs

again become more appealing. In such a scenario, a FRM not only guarantees payment stability but can also be competitive or even less expensive than an ARM.

The study also emphasizes the critical role of prepayment and refinancing options. Prepayment refers to the possibility of paying off part or all the mortgage in advance, before the scheduled due dates, for instance when the borrower has extra funds or wishes to reduce total interest costs. This option offers flexibility in paying off debt faster, particularly under favourable personal financial circumstances. Refinancing, on the other hand, involves replacing an existing mortgage with a new one, usually to take advantage of better terms such as lower interest rates, and it is typically carried out within the first twenty years of the contract. Since refinancing entails fixed costs, its most significant benefits arise when the outstanding debt is still substantial and the potential interest savings are high. Generally, households with greater financial stability, steady incomes, or sufficient savings are more likely to take advantage of refinancing.

Both options are highly relevant: prepayment provides flexibility in managing personal cash flow and interest expenses, while refinancing enables borrowers to respond to changes in broader economic conditions.

The conclusions regarding risk aversion are also confirmed by Coulibaly and Li in their work "Choice of Mortgage Contracts: Evidence from the Survey of Consumer Finances (2007)" which shows that borrowers with higher risk aversion are more inclined to choose FRMs due to the stability of payments. On the other hand, for families that are less risk-averse, the decision between a FRM and an ARM is less influenced by price factors, income volatility and accessibility.

Study 2: "Why do Italian households prefer adjustable-rate mortgages?" by Paola Zocchi, 2013.

The author focuses her analysis on the choices made by Italian households in the residential mortgage market between 1997 and 2006, highlighting that during this period about 75% of the mortgages taken out were ARMs. The study, conducted on a sample of 959 households drawn from the "Survey on Household Income and Wealth" of the Bank of Italy, shows that this strong preference cannot be explained solely by the ability of households to evaluate costs and benefits from an intertemporal perspective. On the contrary, the popularity of ARMs appears to reflect gaps in both financial literacy and risk

perception: rather than considering potential future interest rate increases and the impact of these changes on long-term debt sustainability, many households tend to focus on the short term.

Households demonstrate an even more distorted perspective when their preference for ARMs increases in line with the size of the loan principal and the length of the mortgage. This suggests that, as the loan amount or the repayment horizon grows, households are more likely to choose ARMs, probably because, with longer maturities, the appeal of lower initial payments outweighs concerns about possible future interest rate hikes.

In addition to households' cognitive and informational gaps, Zocchi's study highlights a crucial factor linked to banks' supply policies. In the period under analysis, particularly between 2004 and 2006, banks (driven by strong commercial pressure) progressively lowered their credit standards, favouring ARMs with particularly advantageous initial terms, thanks to the application of temporary discounts (teaser rates). On one hand, these incentives make ARMs significantly more competitive than FRMs, especially for borrowers with higher risk profiles (for example households with low incomes or unstable employment). On the other hand, although this commercial strategy provides apparently advantageous access to credit in the short term, it leads to a less efficient risk allocation in the long run. When rates rise, the increased burden of variable interest payments can become significantly heavier, exposing borrowers to financial distress and increasing the vulnerability of the banking system itself. This dynamic not only undermines borrowers' ability to manage their debt responsibly but also raises questions about the overall stability of the market and the effectiveness of risk distribution among the actors involved. These findings suggest the need for more robust financial education, enhanced informational transparency from intermediaries, and closer oversight of supply policies that favour sales volumes at the expense of long-term stability.

Study 3: "Mortgage Choice in the Euro Area: Macroeconomic Determinants and the Effect of Monetary Policy on Debt Burdens" by Michael Ehrmann and Michael Ziegelmeyer, 2017.

Ehrmann and Ziegelmeyer analyse the euro area using a sample of about 8.700 households to understand how the macroeconomic context and monetary policy decisions influence the preference between FRMs and ARMs. Specifically, the authors show that not only economic growth and labour market stability but also the structure of interest

rates, in the sense of the spread between long-term and short-term rates, and the direction of monetary policy can significantly affect households' borrowing choices, with potential effects on the distribution of well-being among different segments of the population.

In particular, the study highlights that a wide difference between long-term and shortterm rates makes ARMs more attractive, as these are tied to short-term market rates, which are initially lower. This spread prompts households to prefer ARMs over FRMs, reducing the initial cost of the mortgage and making it more convenient to bear future uncertainty related to interest rate fluctuations. At the same time, a context characterized by solid economic growth and low unemployment volatility strengthens households' ability to manage debt. In other words, when expected income is stable or rising, the fear of being unable to deal with possible future interest rate rises decreases, encouraging households to take on the payment variability risk associated with ARMs.

Another relevant aspect concerns differences among households with various income levels. The study shows that wealthier households, thanks to the greater solidity of their balance sheets, are more inclined to choose an ARM under favourable macroeconomic conditions, as they can benefit more from initial lower rates. On the other hand, households with lower incomes, even though they also have access to low interest rates, are unable to fully take advantage of such opportunities, risking the intensification of potential distributional disparities.

Accommodative monetary policies, such as the reduction in short-term rates by central banks, also tend to favour the spread of ARMs. This dynamic has a heterogeneous impact on borrowers: households with higher debt levels and, consequently, higher debt burdens, especially if they already hold an ARM, benefit more from falling rates compared to those with a FRM, which would require a more expensive refinancing process to capitalize on new conditions. As a result, in an expansionary monetary policy environment, more indebted and generally wealthier households might find it even more advantageous to borrow at variable rates, thus widening differences in well-being and inequalities within various segments of the population.

Households are more likely to choose ARMs when economic growth is strong. This suggests that a favourable economic environment encourages households to accept the payment risk associated with ARMs, as they may feel more confident about their future income stability.

Finally, the present text, as well as the previously cited one ("Choice of Mortgage Contracts: Evidence from the Survey of Consumer Finances (2007)"), emphasises how households with a higher level of education show a greater inclination toward ARMs. The premise for this consideration is that a higher level of education corresponds to better financial literacy. The preference for ARMs occurs mainly for two reasons. On one hand, greater financial awareness allows for a better understanding and management of risks, as well as understanding the specific features that distinguish an ARM from a FRM. On the other hand, there is a relationship between a higher level of education and a potential increase in income.

1.3 Gaps and limitations

Despite the rich analytical insights offered by the available literature, the dominant approach in many studies features a series of limitations and simplifying assumptions that diminish models' ability to accurately capture the complexity of household decisions regarding home mortgages.

Firstly, one of the most common hypotheses concerns the complete rationality of households and the availability of perfect information. These premises imply that mortgage decisions are made solely on the basis of strictly economic and rational considerations, in a context free from uncertainty and with complete information. However, reality shows that behavioural factors, cognitive factors (poor financial literacy, errors in probabilistic evaluation), and emotional factors (anxiety or stress related to debt) can significantly influence choices, leading to outcomes different from those predicted by standard models. For example, it is not taken into account that a family might need to purchase a larger or smaller home depending on its needs. Furthermore, the model examines the financial situation resulting from the move but overlooks other inevitably relevant aspects.

In addition, it is often assumed that household savings are invested solely in risk-free assets, ignoring the potentially wide range of financial instruments available, as well as portfolio diversification strategies that might affect mortgage decisions and debt management.

Secondly, the evolution of household income is frequently treated as a stable and predictable factor, with linear growth trajectories. This simplification fails to consider the

impact of unexpected changes in the labour market, fluctuations in earnings, or unforeseen events (such as job loss, illness, or significant shifts in professional careers). Similarly, market conditions are often regarded as exogenous and beyond the predictive capacity of households. In practice, borrowers develop imperfect expectations about future rate trends and may try to forecast potential changes, which in turn influences both when they take out a mortgage and the type they choose. Moreover, the possibility of modifying contractual conditions (such as mortgage renegotiation) or accessing complex financial instruments (such as secondary loans, diversified investments, or risk-hedging products) is sometimes neglected. This omission simplifies the decision-making context, ignoring the dynamic way in which households and intermediaries can adapt their strategies.

From the perspective of credit supply, banks are often considered passive entities that set rates and conditions in response to external factors, without considering their commercial strategies, competitive constraints among intermediaries, or the regulatory framework. This approach overlooks, for example, the active role that lending institutions play in shaping market segments, directing demand toward specific mortgage products, and adopting customized pricing strategies.

Another constraint concerns data quality and availability. Analyses may suffer from issues such as missing data imputation, the lack of key variables (for example, borrowers' expectations about future interest rates, risk perception, or implicit transactional and informational costs), and the inability to fully capture household heterogeneity.

Another important assumption often made, and without which it would be more challenging to build a model, is that the household makes mortgage decisions sequentially, that means deciding whether to choose a fixed or variable rate only after determining the loan amount and duration.

Although these assumptions and simplifications may be useful for constructing a coherent and analytically manageable theoretical framework, they reduce the models' ability to accurately reflect the complex reality of mortgage choices.

1.4 Relation with my analysis

This study focuses exclusively on the binary choice between FRMs and ARMs, excluding other types of mortgages. Furthermore, refinancing cases are not directly considered, as

the corresponding sample would be too small. The reference period is limited to 2001-2010, and it concerns Italian households; the available data derive from a single survey, providing no information on their evolution over time.

The examined literature provides useful insights into the dynamics that lead to the choice between a FRM and an ARM. Specifically, Campbell and Cocco highlight that households with a high degree of risk aversion and uncertain income tend to prefer a FRM, while those with more stable future income may lean towards an ARM, attracted by the lower initial rates. In the analysis conducted, these same variables (or variables related to them) will be incorporated into the model to concretely assess their impact on the sample under consideration. In particular, income stability is also captured by a subjective variable that measures the family's perception of stability, referred to as "Future income expectations". Zocchi addresses this topic as well, noting that an unstable income drives Italian households to opt for an ARM.

Campbell and Cocco also point out that the preference for an ARM in relation to a growing income does not necessarily imply a high initial income; indeed, households facing financial difficulties often aim to reduce costs in the early phases by choosing an ARM. This aspect was examined in the present analysis through the variable *Expenses* compared to income (last 12 months).

Initial amount and duration, both included in the analysis, emerge as central factors in the choice of mortgage type. Zocchi shows that larger amounts and longer durations are associated with a preference for ARMs, since these allow for lower initial payments, while exposing to greater risks in the future. Ehrmann and Ziegelmeyer come to similar conclusion, even though they considered mortgage amount in relation to income.

Another key aspect concerns the impact of monetary policies and the macroeconomic environment, as highlighted by Ehrmann and Ziegelmeyer. These authors indicate that, during periods of robust economic growth and low unemployment volatility, households tend to choose an ARM. In the present analysis the variable *Total real assets* has been included to reflects the household's overall financial solidity.

Finally, the relationship between mortgage choice and financial literacy, as noted by Zocchi, is of primary importance. The gaps in financial education can lead to not-optimal decisions, such as underestimating the long-term risks associated with ARMs. This

phenomenon is also pertinent to the present analysis, in which households with a low level of education are considered less aware of financial literacy.

In the end, the current study aims to compare its findings with those presented in the literature, identifying possible convergences or divergences. By doing so, the study aims to enhance the understanding of mortgage choice dynamics by assessing the extent to which the factors highlighted in previous research are evident in the behaviour of Italian households during the selected period. Ultimately, the research seeks to analyse the actual decisions made by these households and clarify the influences behind them, based on the data available at the time of collection.

Chapter 2 - Methodology & Data

2.1 Data source

The data for the analysis were extracted from the Bank of Italy's website, in particular from "Bilanci delle famiglie italiane" section. They come from an harmonized sample survey on the wealth, income, and consumption of households in the euro area, conducted by the National Central Banks and named the "Household Finance and Consumption Survey (HFCS)". The data collection activities for the first edition, named "Wave 1", were carried out mainly between 2010 and 2011, with a total sample of approximately 62.000 households; for the second edition, named "Wave 2", the interviews involved more than 84.000 households and were conducted mainly between 2013 and 2014.

The documentation used in this study includes both CSV files containing the microdata and the corresponding documentation referring exclusively to the Italian component. In practice, there are files containing a sample of 7.951 households and, respectively, 19.836 individuals for the first wave of the survey, and 8.156 households and, respectively, 19.366 individuals for the second wave.

Starting from the 1989 survey, in order to facilitate the analysis of the evolution of phenomena, a quota of approximately 50% "panel" households was introduced in the sample. These were households already interviewed in previous editions of the survey. The households are selected by ISTAT from the registry lists, and voluntary participation results in a participation rate of just over 50%. The missed interviews represent a problem, as they can lead to samples in which the segments of the population less willing to cooperate are underrepresented; in particular, it has been observed that the difficulty in obtaining the interview increases with the rise in income, wealth, and the education level of the reference person. By definition, the reference person is the head of the household economy or, alternatively, the most informed person among all the members of the household in the year in which the survey is conducted.

2.2 Data selection

The first step was to determine which variables were useful for the analysis. Initially, numerous variables were considered, both those concerning the household's reference individual and those related to the household itself. The variables regarding the individual

categorized him or her from a social, economic, educational, and professional standpoint, while the household-related variables comprised all aspects concerning properties (both the primary residence and others), mortgages, income, assets, and liabilities. The variables initially selected were more than those ultimately used, and part of the objective of this chapter is to illustrate the steps and choices made to arrive at the final analysis. Not only some variables were excluded, either by choice or due to impracticality, but later, others initially neglected, were added.

One of the first decisions regarding data selection was to consider only the first response provided, even when more than one was possible. This choice became necessary for the variables *Occupational status* and *Purpose* in which it was possible to indicate more than one employment condition or more than one reason for taking out the mortgage. This was done in order not to complicate excessively the analysis.

Other decisions came from the scarcity of data or the intent to simplify the study. For example, it was decided to consider exclusively the data related to the first mortgage for each household, since the responses to the question "Did you choose a fixed or variable rate for your second mortgage?" were available in only 20 cases and, similarly, in only 3 cases for the third mortgage. The data lost because of this choice was limited: in wave 1, out of 626 households with at least one mortgage, only 34 had more than one, whereas in wave 2 this phenomenon occurred in 32 out of 644 households.

Another decision was to exclude the case of refinancing, as the data related to this scenario were too few, with 50 observations in wave 1 and 48 in wave 2.

Initially, all data related to "other properties" and not only those concerning the household main residence (HMR) were also considered. The intent was to include a binary variable in the model to indicate whether the type of mortgage (on the HMR or on other properties) influenced the choice between a fixed and a variable rate. This option was in the end not implemented due to the lack of data: in wave 1, for HMR mortgages there were 546 responses out of 626 expected, while for mortgages on other properties only 17 responses were available out of 117 expected; in wave 2, for HMR mortgages there were 623 responses out of 644 expected, while for mortgages on other properties there were 37 responses out of 56 expected.

Regarding the reference years, the complete sample included mortgages obtained from 1981 to 2010 for wave 1 and from 1983 to 2014 for wave 2. This time interval was

considered too extensive for conducting a coherent analysis; therefore, it was decided to limit the analysis to mortgages from 2001 to 2010. The motivations were mainly two: firstly, to exclude the transition period from the lira to the euro; secondly, since the responses provided by households referred to the time at which the survey was conducted rather than the time of the mortgage signing, it was considered necessary to avoid an excessive temporal disparity between the two moments. Furthermore, in wave 2, mortgages after 2010 were eliminated to ensure consistency with wave 1. This choice resulted in a reduction of the sample by approximately 20-30%, with final sizes of 425 for wave 1 and 439 for wave 2, ultimately producing more reliable results.

2.3 Data transformation

After selection, a data transformation process was carried out. It was necessary to delete or set to zero the data that were clearly wrong, this was done without excluding any household or individual. Following this cleaning, the reliability of the remaining data was calculated, defined as the percentage of deleted or set-to-zero values. The formula used was as follows: (Number of transformed values) / (Total number of observations -Number of empty cells) * 100.

In this calculation, the "Total number of observations" corresponds to the total reported for the sample, from which, by subtracting the number of empty cells, the percentage of transformed values relative to the effective number of observations is obtained. Once this percentage was determined, the variables were classified as "reliable" if the percentage was below 5%, "semi-reliable" if between 5% and 15%, "less reliable" if between 15% and 30%, and "unreliable" if above 30%. For derived variables generated through sums or subtractions from other primary variables, the worst category among those of the components was considered.

This strategy allowed, on one hand, to select only more reliable data, but on the other hand, led to the exclusion of some variables that would have been relevant for the regression model. Among these, *Total household gross income* and *Net wealth* would have been particularly indicative in the choice of the type of interest rate, as they would adequately represent the household's wealth. *Net wealth* should have been derived by subtracting the *Total outstanding balance of household's liabilities* from the *Total assets*. The first of these two variables was judged reliable and, therefore, included in the model, whereas *Total assets* was not, as the percentage of values set to zero for one of its

components (*Total financial assets*) was found to be 57% in wave 1 and 51% in wave 2. Regarding *Total household gross income*, the percentage of values set to zero was around 83%, which makes it completely unreliable. Attempts were made to overcome this lack of data by decomposing the variable into its individual components; however, only 30% of the necessary data were found.

Since the two questionnaires were conducted years apart, some of the transformations concerned the compatibility between the wave 1 survey and the one of wave 2. Concerning *Level of education*, wave 1 provided four selectable categories, while wave 2 offered six; to ensure consistency, these categories were reconfigured into four, following the correspondence indicated in the Bank of Italy's documentation, without requiring further assumptions. Regarding *Occupational status*, the filters applied in the various columns of the original file differed; it was adopted the method used in wave 1, which provided for the prevalence of the more general category over the more specific one. For example, if an individual was initially classified as "unemployed", they could not subsequently be indicated as "self-employed". This choice was further motivated by the decision to consider only the first labour status indicated. Moreover, in five responses related to *Occupational status*, it was not specified whether, for employees, the position was permanent or temporary. To overcome this gap and define distinct categories, the variable *Years in main job* was used; since, in these cases, the response was always above six years, a permanent position was attributed.

2.4 Description of variables

In this section, the variables employed in the analysis are described in detail, divided into two main categories. The variables related to the reference person allow the delineation of the individual's demographic and socio-economic profile, while those related to the household provide information on the mortgage, the overall financial situation, the future income expectations and the attitude towards risk. To emphasise the role and relevance of each variable in the model, the following will be reported for each one: the main characteristics, the coding methods employed and any transformations applied.

Variables related to the reference person

The variable *Gender* is a binary qualitative variable that distinguishes between male and female. This variable was not included in the logistic regression model, as it was

considered conceptually irrelevant with respect to the choice between a fixed and a variable rate.

The variable *Age* is quantitative and has been top coded at 85 years. It is approximately referred to the time at which the mortgage was taken out, rather than with the time the interview was conducted. The value was obtained using the formula: Age at mortgage signing = Age in the year of the interview - (Year of the interview - Year of mortgage signing). "Year of the interview" was considered 2011 for wave 1 and 2014 for wave 2.

Age and *Gender* were recorded for every member of the household, whereas the following three variables refer exclusively to individuals aged 16 and over.

The variable *Level of education* is qualitative and represents the highest level of education completed, based on the international ISCED-97 classification used to harmonize educational data internationally. It should be noted that if an individual has not completed a full cycle of studies, the response is coded based on the highest level reached, considering both general and vocational or technical education. Originally, it consisted of four categories (primary or below, lower secondary, upper secondary, tertiary). For simplicity, in the model it was divided into three categories by grouping the first two into a category labelled "basic", alongside "upper secondary" and "tertiary", as the distinction between primary (or below) and lower secondary was considered less relevant.

The variable *Marital status* collects information on marital status, considering both the legal marital status (de jure) and, in some cases, the actual status (de facto). Initially, it included five categories: married, single, divorced, widowed, and consensual union on a legal basis. The latter intended to cover registered unions or similar legal agreements, depending on national regulations. In the Italian context, this category would have included civil unions and cohabitation contracts; however, this became irrelevant as the questionnaire was administered before 2016, the year in which the law on civil unions was introduced. For simplicity, in the model *Marital status* was treated as a binary variable by grouping all categories other than "married" under the label "not married". The aim of this choice was to consider the possibility that a household may have two sources of income rather than one. It should be noted, however, that in some cases this reasoning might not be adequate: indeed, there are married individuals whose household

has only one source of income, while there are unmarried individuals who may receive multiple sources of income, for example, from divorce settlements or spousal pensions.

The variable Occupational status is qualitative and was created by combining three original variables (Labour status, Status in employment, and Type of contract) to initially generate ten categories: employee with a permanent position, employee with a temporary position, self-employed with employees, self-employed without employees, retiree or early retiree, perform domestic tasks, unemployed, student or unpaid intern, unpaid family worker, other. The term "unpaid family worker" refers to someone who contributes to the economic activity of a family member without holding a formal employment contract, living together and actively participating in exchange for a non-monetary form of compensation. Paid interns and priest are included among the employees; moreover, those who work both in their own professional practice and for an employer (public or private) were classified based on the number of hours devoted to each activity. For simplicity, in the model Occupational status was recoded into only three categories: employee, self-employed, and other. This choice was motivated by the intent to evaluate the stability of the reference person's income, assuming that, in the majority of cases, the income of an employee is more stable than that of a self-employed individual, while knowing that in some cases this assumption might not be realistic, for example due to the presence of temporary contracts among employees.

Variables related to the household

The variable *HMR tenure status* is qualitative and comprises four categories: own all, own part, rented/sublet, and free use (which include also usufruct). This variable is binding for the subsequent sections of the questionnaire, as only households belonging to the first two categories could answer the questions aimed at identifying those who have a mortgage. The variable was not included in the model since, as will be shown in the descriptive statistics, the "own part" category represents only about 0.5% of the sample, while almost all cases fall into "own all", making it irrelevant for the analysis.

The variables *Mortgages using HMR as collateral* and *Mortgages using other properties as collateral* have been defined as binary variables, with responses "yes" or "no", and aim to detect how many households have actually taken out a mortgage, distinguishing between those who use their own HMR and those who use other properties as collateral. It is important to recall that collateral represents an asset, or a set of assets, pledged as

security for the repayment of the loan, for example, a house, a car, a property, or machinery. In the event of the borrower's default, the pledged assets will be seized and sold, and the proceeds used for loan repayment. These variables would have been included in the model to study the effect of the type of collateral on the choice of interest rate, but as already mentioned, this was not possible due to the lack of data related to mortgages on other properties. Consequently, in the chosen subset only those mortgages for which the collateral is the HMR were considered.

The variable *Purpose* is qualitative and, for simplicity, was included in the model as a binary variable, distinguishing between mortgages taken out for the purpose of purchasing the HMR and those with other purposes. In the latter case, various objectives are comprised, such as purchasing another real estate asset, renovating or renewing the residence, buying a vehicle or other means of transport, financing an entrepreneurial or professional activity, consolidating other consumer debts (that is, the financial operation that allows the unification of multiple pre-existing loans or debts to obtain more favourable conditions), and covering living expenses or other purchases.

The variable *Year of signing* was transformed into a categorical variable for the model by dividing the years of mortgage signing into the following groups: 2001-2002, 2003-2005, 2006-2007, and 2008-2010.

The variable *Initial amount* is quantitative and indicates, in euros, the amount initially borrowed for the mortgage, while the variable *Duration* represents, in years, the length of the mortgage.

The variable *Current value all properties*, also quantitative in euros, represents the household's real estate wealth and was obtained by summing: the current value of the HMR, the current value of other properties, and the additional properties' current value. The current value of the HMR was calculated using the formula: (% ownership HMR / 100) * Current price of the HMR. The current value of other properties was obtained by summing the values of up to three properties, calculated in a similar manner. The third addend refers instead to the value of any additional properties not included in the "other properties" category.

The variable *Total real assets* is quantitative in euros and is obtained by summing: the value of the HMR, the value of other real estate, the value of vehicles owned by the household, any valuable objects and the value of assets derived from self-employment.

Although it was necessary to set some values to zero, the variable was used in the model since the percentage calculated was about 0.95% for wave 1 and 1.35% for wave 2, so it has been considered reliable. In any case, to improve the estimate, it was verified that the indicated value was at least equal to that of *Current value all properties*; if this was not the case, the value was manually fixed.

The variable *Outstanding balance household's liabilities*, quantitative in euros, indicates the total amount of outstanding debts to be repaid, including not only the mortgage but also other financing, personal loans, and similar. It is composed of two main components. The first represents the outstanding balance of mortgage debt (the "collateralised liabilities" or "secured debt") related to all properties owned by the household, obtained by summing the mortgage debt secured by the HMR and that secured on other properties. The second comprises the outstanding balance of other non-mortgage debts (total unsecured debt or non-collateralised liabilities), which includes balances related to lines of credit, bank overdrafts, interest-bearing credit cards, and all other loans (for example car loans, consumer loans, instalment payments, and private loans from relatives, friends, or employers).

The variable *Expenses compared to income (last 12 months)* is qualitative and derives from the question in the survey: "Aside from any purchases of assets, over the last 12 months would you say that your (household's) regular expenses were higher than your (household's) income, just about the same as your (household's) income, or that (you/your household) spent less than (your/its) income?" The three categories that compose it are obviously: expenses higher than income, expenses similar to income, and expenses lower than income.

The variable *Future income expectations*, also qualitative, originates from the question: "Over the next year, do you expect your (household's) total income to go up more than prices, less than prices, or about the same as prices?" And the three respective categories compose it.

Finally, the variable *Risk attitude* is qualitative and arises from the question: "Which of the following statements comes closest to describing the amount of financial risk that you (and your husband/wife/partner) are willing to take when you save or make investments?" The response options are four: "take substantial financial risks expecting to earn substantial returns", "take above average financial risks expecting to earn above average

returns", "take average financial risks expecting to earn average returns" and "not willing to take any financial risk".

2.5 Descriptive statistics

In this section, some graphs related to the descriptive statistics are presented and discussed, further details are available in the appendix A. It should be recalled that the chosen subset refers exclusively to households that answered the question regarding the choice of the interest rate (fixed or variable) and that took out a mortgage in the period between 2001 and 2010, also considering only the reference person.

Descriptive statistics related to the reference person

The variable *Age* shows that, although the initial number of observations is greater in wave 1, after sample selection the number is slightly higher in wave 2. This may be due either to the overall greater number of households in wave 2, which would imply that households in wave 1 include a larger number of members in addition to the reference person, or to a slightly higher percentage of available data in wave 2. The mean and median indicate that individuals in wave 1 are, on average, slightly younger. Moreover, the standard deviation of the chosen subset is lower, as expected, since individuals under 26 years of age were excluded from this subset, this ensured a more homogeneous distribution.

The variable *Gender*, in the full sample, presents a balanced distribution with 52% women in both waves; however, in the chosen subset the percentage of women drops significantly at 37% in wave 1 and at 42% in wave 2. Since this difference is not observed in the full sample, it is hypothesized that within each household the reference person is more frequently selected as male, probably due to income differences or because men are statistically more informed about household finances.

For the subsequent variables, the number of observations in the full sample is around 17.200, given that the questionnaire considers only individuals aged 16 and over, while no missing data are found in the chosen subset, in line with expectations.

The analysis of the variable *Marital status*, supported by the exclusion of individuals under 26 years of age, reveals a higher percentage of married individuals in the chosen subset. Although the "not married" category also includes widowed and divorced people,

the overall effect tends to favour the married, since the combined percentages of divorced and widowed individuals are generally lower than that of single individuals. Moreover, the fact that the chosen subset considers exclusively households that have taken out a mortgage may have further shifted the distribution toward the "married" category, as married individuals are more likely to apply for a mortgage. The differences observed between the two waves, which could suggest a temporal trend toward a lower marriage rate in wave 2, are also evident in the subset, indicating that this effect is attributable both to temporal factors and to the sample structure.



The analysis of *Level of education* shows that, since the chosen subset is composed of individuals older than 26 years, the percentage of people in the "basic" category is lower, while the proportions of those with "upper secondary" and "tertiary" education increase. This result can be explained both by the older age of the respondents and by the fact that, from the 2000s onward, greater importance was attributed to education compared to the 1980s, resulting in a shift in the predominant category from "basic" to "upper secondary". Slight differences between the two waves are observed in the full sample; the subset maintains consistency between wave 1 and wave 2.



Finally, the analysis of *Occupational status* reveals that, in the chosen subset, the majority of reference people are classified as either employee or self-employed, with these

categories being doubled compared to the full sample. In particular, the employee category is considerably more significant, suggesting that within the household the reference person is more frequently an employee, likely due both to the larger number of employees and to the fact that banks tend to regard employment income as more secure and verifiable. On the contrary, the full sample shows a greater incidence of the "other" category, which includes retirees, students, unemployed individuals, and those performing domestic tasks. It is also observed that in wave 2 the percentage of employees is slightly lower and that of "other" is higher, both in the full sample and in the chosen subset, which, although altering the expected occupational distribution, still maintains considerable consistency between the two waves.



Descriptive statistics related to the household - Quantitative variables

For the variables *Current value all properties* and *Total real assets*, the number of observations is consistent with the available data and with the choice of the subset, there are no missing data. The maximum values of both variables are significantly higher (almost double) in wave 1, as also reflected in the median and the mean, whose difference amounts to approximately 12% of the average value between the two (28.000/230.000 = 0.12). One possible explanation is that the survey period for wave 2 is slightly shifted relative to wave 1, including the years of the sovereign debt crisis (from 2010 to 2014), a period during which rising interest rates, increased unemployment, and higher tax pressure led to a general decline in property values and, consequently, in real assets. In the selected sample, this condition is accentuated, with a difference between the means amounting to approximately 27% of the average value between the two (85.000/315.000 = 0.17). Furthermore, while in the full sample the minimum values may be zero, in the chosen subset they assume values greater than zero, implying that households without

properties or real assets do not obtain a mortgage, which is consistence with its definition. It is also interesting to note that the maximum values are much lower in the subset, suggesting that households particularly wealthy in terms of properties and real assets are probably not mortgage holders. At the same time, outliers may be present in the full sample, as supported by the considerable difference observed between the maximum value and the mean.

For the variable Outstanding balance of household's liabilities, the number of observations is consistent with the available data, there are no missing data. The maximum value of the liabilities is higher in wave 1, and this difference is reflected in a mean that is approximately 29% higher compared to wave 2. This result is consistent with the fact that, in the presence of greater real assets, households tend to incur higher liabilities, such as larger mortgages. Similarly, it is plausible that the sovereign debt crisis induced a reduction in indebtedness, either as a precaution or due to fewer credit opportunities, as evidenced by the fact that, in the subset, the mean in wave 1 is approximately 28% higher. It is interesting to note that, in a sample of households with a mortgage, the outstanding mortgage balance is not expected to be zero (except in cases of recent payoff). In this context, the fact that the minimum is zero only in wave 2 means that, by chance (following the logic of probabilities, as a higher percentage of zeroed values makes the observation of a zero more likely), in the process of sample selection in wave 1 no zero values were selected, whereas in wave 2 this happened. Note that in wave 2 the variable was classified as semi-reliable (about 10% of zeroed values), whereas in wave 1 it is considered reliable (about 3%).

Regarding *Initial amount*, the number of observations corresponds to the number of mortgages, and there are no missing data. The minimum, as expected, is never zero, since taking out a mortgage always involves an initial amount. Although the maximum value is higher in wave 1 (probably due to an outlier), the mean in wave 2 is approximately 7.5% higher. At first sight, the fact that the average amount requested in wave 1 is lower compared to wave 2 appears counterintuitive, considering that in wave 1 the real assets are greater and, theoretically, households could afford larger mortgages. However, it is plausible that households with greater patrimonial assets may have self-financed part of the property purchase, thereby reducing the mortgage amount, whereas households with fewer real assets in wave 2 may be forced to finance a larger share of the price. Another plausible explanation is related to post-crisis banking policies, which may have favoured

a self-selection process in which only those who can demonstrate sufficient income obtain a mortgage and, once approved, request a higher amount. In fact, the crisis does not necessarily reduce the mortgage amount but rather limits access to it. Although the average mortgage amount in wave 1 is lower than in wave 2, which appears counterintuitive even considering that wave 2 includes the years of the sovereign debt crisis, during which rising interest rates should have discouraged large mortgages, it should be noted that the official central bank rates during that period (2012-2013) were lowered as a countermeasure. Moreover, some households in wave 2 may still have requested substantial amounts, either because house prices did not drop in all areas or because a high mortgage was necessary to make a purchase. This condition, however, is not particularly reflected in the selected sample, where the means of the two periods are practically identical. In other words, in a more restricted subset (mortgages from 2001 to 2010) the difference disappears, suggesting that the high-value mortgages in wave 2 may pertain mainly to periods outside the considered interval, that is, after 2010 or before 2001.

Finally, the variable *Duration* presents some gaps due to missing responses, which are negligible: 7 observations out of 626 in wave 1 and 2 out of 644 in wave 2. The average duration of the mortgages is slightly longer in wave 2 compared to wave 1, a difference that is also observed in the subset. Since this difference persists in the chosen subset, it is unlikely to be solely due to temporal factors; a possible explanation is that larger mortgages tend to have longer durations, this hypothesis will be further verified through the analysis of the correlation matrix.

Descriptive statistics related to the household - Qualitative variables

For the variable *HMR tenure status*, the number of observations is consistent with the available data and with the choice of the subset, so no missing data are present. In the full sample, the majority (around 70%) of households fully own the HMR (this might be due to cultural and historical factors), while a significant percentage is represented by households that rent or sublet it; only around 1-2% of households partially own the HMR. Consequently, in the selected sample, 98-99% of the observations fall into the "own all" category, making the variable irrelevant for the model, which is why it was not considered. The differences between wave 1 and wave 2 are negligible.

Regarding mortgages, the variable concerning the use of the HMR (or other properties) as collateral confirms, as previously stated, that only mortgages on the HMR were considered in the analysis. The number of observations and the percentage of affirmative responses correctly reflect the number of mortgages present in the full sample.

Analysing the variable *Purpose*, the number of observations is in line with the number of mortgages and with the subset selection, with no missing data detected. As expected, the majority of mortgages were taken out for the purpose of purchasing the HMR, in fact grouping together the other purposes the percentage obtained is about 25%. In particular, the prevalence of mortgages taken out to purchase the HMR is even more pronounced in wave 2 compared to wave 1, with a difference of approximately 7 percentage points. This difference could be explained by the crisis of the years 2010-2014, during which a reduction in wealth led to a greater necessity to obtain a mortgage for the purchase of the HMR, while other purposes (for example non-urgent renovations) tend to be postponed. In the selected subset, the predominance of the purpose "to purchase the HMR" is even more marked, although the difference between the two waves is reduced.

The variable *Year of signing*, in the full sample, shows a number of observations consistent with the number of mortgages in wave 2, while in wave 1 there are 21 missing data. In the chosen subset no missing data are observed. The selection excludes the categories "before 2001" and "after 2010". The exclusion of these categories, which had different weights in the full sample, affects the distribution of the variable in the chosen subset. Therefore, the differences between the waves are more accurate in the subset, where, for two of the four considered periods, there are no differences.



The variable *Expenses compared to income (last 12 months)* shows a number of observations consistent with the available data, with no evidence of missing values. The

differences between the graphs of the full sample and the subset are minimal, suggesting that the selected sample accurately represents the general sample. A small increase in the percentage of cases where expenses exceed income is observed when considering only households with a mortgage. This phenomenon appears counterintuitive, although the monthly mortgage payments constitute an expense that may contribute to this outcome. A big difference emerges between the two waves: in wave 2 the predominant response is "expenses lower than income", whereas in wave 1 the most frequent response is "expenses about the same as income". In the full sample, the majority of households in wave 2 report expenses lower than income, probably due to the more stringent post-crisis mortgage conditions; however, the same trend is also found in the subset, suggesting that this difference may depend also on intrinsic financial behaviours rather than exclusively on the period considered.



For the variable *Future income expectations* missing data in the full sample are around 7%, while in the chosen subset they are reduced to approximately 4%. In the subset, an increase in the percentage of responses indicating "less than prices" is observed, accompanied by a decrease in those indicating "more than prices", suggesting greater pessimism among households with a mortgage compared to the general sample. It is clear, therefore, that in wave 2 pessimists and optimists are balanced, whereas in wave 1 prevails pessimism. The same observations apply to both the full sample and the chosen subset, indicating that the phenomenon cannot be attributed only to years outside the considered range. A plausible explanation for this difference is the larger number of open mortgages between 2001 and 2002 in wave 1 compared to wave 2 (approximately 40 more mortgages). In those years, the general sentiment in Italy tended to be more cautious and concerned due to uncertainties related to the change from physical currency, the

perceived purchasing power, and the international economy shaken by the tragic event of 2001. Other reasons may be intrinsic to the composition of the sample.



Finally, the variable *Risk attitude* presents a number of observations consistent with the original data and with the chosen subset, with no missing values. Comparing the graphs, a slight increase in the percentage of cases classified as "moderate risk taker" is observed in the subset, accompanied by a slight reduction in the "risk-averse" cases. This shift suggests that some households tend to move from the "risk-averse" category to "moderate risk taker", while remaining overall oriented towards a low propensity for risk. It is interesting to note that, although households in wave 1 appear more pessimistic regarding the future increase in income relative to inflation, they are also willing to take greater risks to obtain higher returns. Such behaviour may derive either from a conscious choice aimed at compensating for negative income prospects with riskier investments, or from a possible overestimation of opportunities. Moreover, the fact that wave 1 is composed of younger individuals may contribute to a higher propensity for risk; this will be partially confirmed by the analysis of the correlation matrix.



Chapter 3 - Results analysis

3.1 Correlation matrix

The correlation matrices (reported in appendix B) were constructed by eliminating the values below the diagonal, as these would be equal to those above it, thereby obtaining upper triangular matrices. On the diagonal, the correlation of each variable with itself is present, which is by definition equal to 1, while the other values can range from -1 to approximately 1.

In the case of correlations between continuous variables, a value greater than zero indicates a positive correlation (as one variable increases, the other increases proportionally), whereas a value less than zero indicates a negative correlation (as one increases, the other tends to decrease). In addition to the correlations between continuous variables, correlations between a continuous variable and a categorical one were also analysed. In this latter case, it is observed that when the category in question is associated with a positive correlation, the average value of the continuous variable tends to be higher, and vice versa in the presence of a negative correlation. When there are more than two categories, it should be considered that the other categories are treated as a whole.

Below there is an analysis of the individual correlations common to both waves, in order to exclude fortuitous cases and focus only on those evident in both samples; the obvious correlations between the categories of the same variable are not discussed.

It is observed that the correlation between *Current value all properties* and *Total real assets* is very strong and positive in both waves, with coefficients above 0.9. This high correlation derives from the intrinsic nature of the variables, as a family's total real assets largely consist of the value of its properties.

The correlation between *Initial amount* and *Outstanding balance of household's liabilities* is strong, particularly in wave 1, where the coefficient exceeds 0.7. This result is consistent with the fact that, when a household takes out a mortgage, a large part of the associated liabilities is represented by the amount itself. The lower correlation value in wave 2 suggests that, in this second sample, liabilities arising from loans not related to the purchase of the home constitute a larger share of the overall liabilities. This result is also influenced by the presence of some zeroed values in wave 2 that remained in the sample (as noted in the descriptive statistics).

The correlation between *Initial amount* and *Duration* is moderately positive, confirming what was observed in the descriptive statistics: larger mortgage amounts are associated with longer durations.

The correlation between *Initial amount* and *Total real assets* is positive, weak in wave 1 and moderate in wave 2. This correlation indicates, as expected, that households possessing greater tangible assets can afford larger mortgages, since banks are more inclined to grant higher amounts when the household has solid assets.

A weak, negative correlation is observed between *Duration* and *Age*. A plausible explanation is that for older individuals, the number of years remaining before retirement is lower, which pushes or forces them (if the request comes from the bank) to choose a shorter mortgage duration.

The correlation between *Initial amount* and *Purpose* shows that, in mortgages taken out for the purpose of purchasing the HMR, the average amount requested is higher compared to mortgages with other purposes. This finding is predictable, considering that purposes not involving the purchase of the HMR usually require lower amounts.

Similarly, the correlation between *Duration* and *Purpose* indicates that for mortgages aimed at purchasing the HMR, the average duration is longer. This relationship is consistent with the observation that *Initial amount* and *Duration* are positively correlated.

It is also noted that as the level of education increases, the average wealth of the household increases, as evidenced by the correlation between *Total real assets* and *Level of education*. Although the values are not particularly high, an increase in the correlation is observed when moving from a lower to a higher level of education. The correlations are more marked in wave 1, where households on average are wealthier. This evidence suggests on one hand that a higher level of education may lead to a more comfortable economic situation, given the higher expected income, and on the other hand that already wealthy families may invest more in education.

Finally, the correlation between *Level of education* and *Occupational status* reveals that, in the presence of a basic education, it is more likely for the person to belong to the "other" category for occupational status, rather than falling into the employee or self-employed categories. This phenomenon is partly due to the fact that the "other" category includes, among others, the unemployed, who might not have achieved a sufficient level of education to obtain employment or start their own business. Cases related to students, unpaid interns, or unpaid family workers were not considered, as these categories are not present in the selected sample. On the contrary, in the presence of a "upper secondary" education the only positive correlation found is with the employee category, whereas for a "tertiary" education the correlations are positive with both the employee and the self-employed categories.

3.2 Logistic regression

Multicollinearity

After examining the correlation matrix, it was appropriate to limit the effects of multicollinearity by excluding from the model one of the two variables for each strong correlation. The variable *Current value all properties* was excluded, as it is closely correlated with *Total real assets*, and the variable *Outstanding balance of household's liabilities* was excluded, being correlated with *Initial amount*. The choice between the first two was based on the fact that *Total real assets* is more comprehensive than only the value of the properties, while among the latter, it was decided to retain *Initial amount*, since it is frequently cited in the literature as a parameter that influences the decision.

Theoretical background

In the logistic regression model, the "Value" represents the estimated coefficient, which allow to understand the direct relationship between the explanatory variable (predictor) and the outcome. In practice, the coefficient indicates how much the logit (or log-odds) of the outcome changes for a one-unit increase in the independent variable, holding all other variables constant. The adopted formula is: logit (p) = log (p/(1-p)). A positive coefficient implies that an increase in the independent variable leads to an increase in the probability of the event, which is, the choice of the variable interest rate. When a qualitative variable has k categories, the model includes only k-1 indicators, excluding one category that serves as the reference, in order to avoid issues of multicollinearity. A positive "Value" indicates that belonging to the category in question increases the probability of the event with respect to the reference category.

In the case of standardized parameters, each independent variable is first transformed by subtracting its mean and dividing by its standard deviation, thereby placing all variables on the same scale. The standardized coefficient indicates how much the logit of the outcome changes for an increase of one standard deviation in the variable, holding all others constant. This approach permits an assessment of the relative importance of each variable, making it evident which one has the strongest effect on the log-odds of the outcome. This becomes particularly interesting in models characterized by heterogeneous units of measurement, such as euros, years, and units.

The "Standard error" represents the measure of variability of the estimated coefficient and indicates how much the coefficient might vary if the model will be repeated on different samples. A small standard error suggests a more precise estimate of the coefficient.

The "Odds Ratio (OR)" is calculated as the exponent of the coefficient and represents the change in the ratio between the probability of "success" (choosing the variable interest rate) and that of "failure" (choosing the fixed rate) for a one-unit increase in the independent variable. The odds ratio is defined as: p/(1-p).

The p-value, indicated as " $Pr > Chi^{2}$ ", derives from the Wald test, which evaluates the null hypothesis that the coefficient of an independent variable is equal to zero, that means that the variable does not affect the outcome. A p-value lower than the chosen significance threshold implies that the coefficient (and consequently the odds ratio) is significantly different from zero (or from 1, in the case of the OR), suggesting a real effect of the variable. Conversely, a high p-value indicates the impossibility to exclude that the coefficient is the result of chance.

Finally, the significance threshold (alpha) represents the critical level adopted to determine whether an effect is statistically relevant. An alpha value of 0.05 (5%) is commonly used, while a level of 0.10 (10%), although considered less rigorous, is sometimes adopted in exploratory analyses in order not to exclude potential signals of effect. A level of 0.20 (20%) is much less common, as it increases the risk of false positives, that is, affirming the existence of effects that may not actually be present. In general, raising the value of alpha makes the test more permissive, while lowering it increases the risk of false negatives (an effect may not be detected due to the strict threshold).

Model parameters

In this section, the estimated coefficients from the logistic regression model are examined, detailing the meaning and interpretation of each parameter. In the tables presented below, green corresponds to a 5% threshold, blue to 10%, and yellow to 20%.

Source	Value	Standard error	Pr > Chi²	Odds ratio
Intercept	1,879	0,979	0,055	
Initial amount	-0,005	0,002	0,014	0,995
Duration	-0,010	0,019	0,606	0,990
Total real assets	0,000	0,000	0,946	1,000
Age	-0,003	0,013	0,790	0,997
Purpose-1	0,032	0,293	0,914	1,032
Purpose-2	0,000	0,000		
Year of signing-1	-0,866	0,323	0,007	0,421
Year of signing-2	-0,965	0,284	0,001	0,381
Year of signing-3	0,358	0,331	0,280	1,431
Year of signing-4	0,000	0,000		
Expenses compared to income-1	-0,252	0,360	0,484	0,777
Expenses compared to income-2	-0,101	0,245	0,679	0,904
Expenses compared to income-3	0,000	0,000		
Future income expectations-1	-0,504	0,431	0,241	0,604
Future income expectations-2	0,096	0,246	0,697	1,100
Future income expectations-3	0,000	0,000		
Risk attitude-1	-0,118	0,801	0,883	0,889
Risk attitude-2	-0,815	0,309	0,008	0,443
Risk attitude-3	-0,359	0,243	0,140	0,699
Risk attitude-4	0,000	0,000		
Marital status-1	-0,017	0,284	0,952	0,983
Marital status-2	0,000	0,000		
Level of education-1	0,257	0,325	0,429	1,293
Level of education-2	0,002	0,285	0,994	1,002
Level of education-3	0,000	0,000		
Occupational status-1	-0,423	0,310	0,172	0,655
Occupational status-2	-0,466	0,384	0,225	0,627
Occupational status-3	0,000	0,000		

Fig.1 - Model parameters wave 1 (Variable interest rate as the reference category)

Source	Value	Standard error	Pr > Chi ²	Odds ratio
Intercept	1,376	1,039	0,185	
Initial amount	-0,003	0,002	0,176	0,997
Duration	-0,043	0,018	0,016	0,958
Total real assets	0,001	0,001	0,047	1,001
Age	-0,010	0,012	0,421	0,990
Purpose-1	0,483	0,303	0,110	1,621
Purpose-2	0,000	0,000		
Year of signing-1	-0,057	0,365	0,875	0,944
Year of signing-2	-0,530	0,273	0,052	0,589
Year of signing-3	0,185	0,280	0,509	1,203
Year of signing-4	0,000	0,000		
Expenses compared to income-1	0,271	0,321	0,398	1,311
Expenses compared to income-2	0,305	0,255	0,232	1,357
Expenses compared to income-3	0,000	0,000		
Future income expectations-1	-0,665	0,483	0,168	0,514
Future income expectations-2	-0,325	0,471	0,490	0,723
Future income expectations-3	0,000	0,000		
Risk attitude-1	-1,184	1,177	0,314	0,306
Risk attitude-2	-0,403	0,381	0,290	0,668
Risk attitude-3	-0,020	0,233	0,930	0,980
Risk attitude-4	0,000	0,000		
Marital status-1	0,137	0,249	0,581	1,147
Marital status-2	0,000	0,000		
Level of education-1	0,022	0,320	0,945	1,023
Level of education-2	-0,258	0,278	0,352	0,772
Level of education-3	0,000	0,000		
Occupational status-1	-0,086	0,309	0,780	0,917
Occupational status-2	-0,602	0,404	0,136	0,548
Occupational status-3	0,000	0,000		

Fig.2 - Model parameters wave 2 (Variable interest rate as the reference category)

Intercept

The intercept in a logistic regression represents the natural logarithm of the odds for the reference class (in this case, the variable interest rate) when all independent variables are equal to zero. When the model's variables are not centred on zero (their mean is not close to zero), the intercept indicates a probability based on an unrepresentative scenario, deviating from the overall sample probability, which is approximately 50% for wave 1 and 55% for wave 2. The predicted probability from the intercept can be calculated using the formula: P (adjustable when all independent variables equal 0) = $e^{(intercept)}/(1 + e^{(intercept))}$.

With an intercept of 1.879, this yields a probability of approximately 87% in wave 1, while in wave 2 the same procedure results in a probability of about 80%.

Initial amount

The variable *Initial amount* has been expressed in thousands of euros to avoid that, using the euro as the unit, an increase of one euro produces a coefficient so small that it appears zero when rounded. This transformation makes the coefficient clearer to be interpreted. Although the coefficient is not very significant in wave 2 (falling only within the 20% threshold), it is highly significant in wave 1 (nearly at the 1% threshold). In both waves, the coefficient is slightly negative, implying that as the initial amount increases, the probability of choosing a variable rate decreases. In particular, an odds ratio of about 0.996 indicates that for every increase of 1000 euros the probability of opting for a variable rate decreases by roughly 0.4%.

This result is consistent with the considerations of Campbell and Cocco, who claim that, especially for those with uncertain or highly volatile income sources, as the amount borrowed increases it is preferable to choose a fixed rate in order to avoid fluctuations in the monthly payments. A similar conclusion was reached by Ehrmann and Ziegelmeyer, who relate the initial amount to income; in contrast, Zocchi's work argues that with a higher initial amount, households tend to prefer an ARM to facilitate lower initial payments. However, this bias does not appear to be present in the two analysed samples, although the small coefficient does not rule out that some households might be influenced by this effect. In general, the choice of a FRM in the presence of a high amount appears consistent, so households opting for a variable rate are likely characterized by a greater interest in short-term savings.

Duration

Regarding the variable *Duration*, the coefficient is not significant in wave 1, whereas in wave 2 it is highly significant, almost reaching the 1% threshold. In wave 2, the coefficient is slightly negative, indicating that as the mortgage duration increases, the probability of choosing a variable rate decreases; specifically, an odds ratio of about 0.96 implies that for each additional year, the probability of opting for a variable rate decreases by roughly 4%.

This result is consistent with the observed correlation between *Initial amount* and *Duration* and is logical considering that the exposure to a variable rate for long periods entails a high risk of fluctuations. However, this outcome contrasts with the theories of Campbell & Cocco and Zocchi, which suggest that a longer duration would be associated with a greater tendency toward an ARM, both to achieve lower initial payments and for the potential in long-term savings.

Total real assets

The variable *Total real assets* was also expressed in thousands of euros to make the coefficient more interpretable. This transformation has been effective only in wave 2, where the coefficient is non-zero and significant (within the 5% threshold), whereas in wave 1 the coefficient is zero and not significant.

In wave 2, a slightly positive coefficient implies that as the family's real asset value increases, the probability of opting for a variable rate also increases; in particular, an odds ratio of about 1.001 indicates that for every increase of 1000 euros, the probability of choosing a variable rate rises by approximately 0.1%. This result is consistent with the idea that greater financial stability, as indicated by higher real assets, may reduce concerns over potential rate hikes, thereby enabling households to take advantage of the initial benefit offered by an ARM, as evidenced by Ehrmann and Ziegelmeyer.

Age

Regarding the variable *Age*, the coefficients are not significant in either sample. Probably the impact has been overshadowed in the model by the stronger effect of other variables. It was expected that younger individuals, who anticipate higher income growth and show greater sensitivity to initial expenditures, would prefer a variable rate.

Purpose

Regarding the variable *Purpose*, the coefficient is not significant in wave 1, while in wave 2 it is weakly significant, nearly reaching the 10% threshold. In wave 2, the coefficient associated with the purpose "To purchase the HMR" is positive, which implies that when the purpose is the HMR the probability of choosing a variable rate increases compared to when the purpose is different. In particular, an odds ratio of approximately 1.62 suggests that all other conditions being equal, the probability of opting for a variable rate increases by 62% when the purpose is the purchase of the HMR compared to the reference category "other".

This result contrasts with the findings of Ehrmann and Ziegelmeyer, who claims that, in the case of a mortgage to purchase the HMR, the preferred rate is fixed because, with other purposes, risk aversion tends to decrease. However, this effect was not observed either in the correlation matrix or in the regression. In general, it can be stated that the choice of a FRM intended for the purchase of the HMR is plausible, given that both the amount and the duration tend to be high; households that opt for a variable rate are likely characterized by a greater interest in saving on short-term payments.

Year of signing

Regarding the variable *Year of signing*, it is observed that in wave 1 two of the four coefficients related to the considered periods are highly significant (both falling within the 1% threshold), whereas in wave 2 only one coefficient is significant, nearly reaching the 5% threshold. In wave 1, the coefficient for the years 2001-2002 is strongly negative, which implies that when the mortgage was taken out during this period, the probability of choosing a variable rate is lower compared to mortgages signed in the period 2008-2010. In practical terms, an odds ratio of about 0.42 indicates that, all other conditions being equal, the probability of choosing a variable rate decreases by 58% when the mortgage is taken in the 2001-2002 period, relative to the reference category (2008-2010). Similarly, in both waves the coefficient for the period 2003-2005 is strongly negative, with an average odds ratio of approximately 0.48, which suggests a 52% decrease in the probability of choosing a variable rate for mortgages signed during this period compared to those from 2008-2010.

Before examining how these results compare with both expectations and the literature, a brief summary of the inflation trend during the considered period and its relationship with

the official interest rate is provided. The official interest rate is the rate that banks pay on funds borrowed from the Central Bank, and it influences the level of all other interest rates. This link arises from the fact that lenders are interested in obtaining a positive (or at least non-negative) real rate; consequently, in the presence of rising inflation, banks and lenders demand higher rates to ensure a real gain. In response to inflationary growth, the Central Bank tends to increase the official rate, so that interest rates follow the inflation rate, unavoidably with some delay.

The introduction of the euro in the early 2000s contributed to greater inflation stability and a favourable economic environment, characterized by low interest rates and improved access to credit. During the period 2001-2007, inflation remained stable at around 2%, in line with the ECB's target, and this stability was one of the main reasons for the reduction in official rates. Just before the 2008 crisis, there was a slight increase in inflation, while during the crisis (2008-2009) inflation dropped drastically, pushing the ECB to adopt an expansionary monetary policy that led to a strong reduction in the official rate.

Considering the individual periods, the period 2001-2002 was characterized by an economic slowdown starting from the late 1990s, aggravated by the tragedy of the Twin Towers at the end of 2001, which increased market uncertainty and led the ECB to lower rates from around 4% to 2%. In the period 2003-2005, despite a slight economic recovery, geopolitical uncertainties persisted (for example, the outbreak of the Iraq war), pushing the ECB to keep interest rates low, below 3%. In the period 2006-2007, an economic recovery in Europe coupled with rising commodity prices generated concerns about inflation, leading the ECB to raise rates back to around 4%. Finally, in the period 2008-2010, following the global financial crisis in September 2008 (marked by the collapse of Lehman Brothers) and the subsequent recession, the ECB drastically reduced rates, lowering them from 4% to 1%, reaching an historic minimum.

These economic contexts can explain households' preferences: during 2001-2005, uncertainty and concern for the future, along with favourable fixed rates compared to previous years, pushed households to choose FRMs; whereas during 2008-2010, an extremely low official rate, combined with a general lack of awareness that this rate was derived from ECB policy (which could have subsequently increased once the crisis passed), favoured the choice of a variable rate. These preferences are consistent with the observations reported by Ehrmann and Ziegelmeyer, according to whom, if a household

anticipates a future reduction in rates, the choice falls on an ARM, while in the presence of stable or rising rates a FRM is preferred.

Expenses compared to income (last 12 months)

For the variable *Expenses compared to income (last 12 months)*, the coefficients are not significant. For example, considering a household with expenses exceeding income, where cash flows are limited, one might hypothesize that it would prefer an ARM if the goal is to spend less immediately, or a FRM if there is a fear of being unable to meet unforeseen expenses. However, the lack of significance suggests that, in this model, the impact of this variable has been obscured by the influence of other factors.

Future income expectations

Moving on to the variable *Future income expectations*, in wave 1 the coefficients are not significant, whereas in wave 2 one of the two coefficients is weakly significant (falling within the 20% threshold). The weakly significant coefficient refers to the category "more than prices" and is strongly negative, which implies that when a household expects its income to grow more than prices, it tends to choose a variable rate less frequently than households that expect their income to grow roughly in line with prices. In practical terms, an odds ratio of approximately 0.51 indicates that, all other conditions being equal, when the income is expected to grow more than prices the probability of opting for a variable rate decreases by 49% compared to when it is expected to grow about the same as prices.

Although this result is only weakly significant, it is contrary to the expectations and conclusions of Campbell and Cocco, who argue that income growth exceeding inflation would indicate a genuine enrichment and the ability to bear potential rate increases, whereas if income grows at the rate of inflation, the safer choice would be a FRM. Campbell and Cocco take into account the impact of income uncertainty, contending that when income is uncertain the most appropriate choice is a FRM, while in the presence of stable or growing income (even in economically challenging situations) the best choice would be an ARM. A possible explanation for the contrary result obtained is that households' expectations of income growth are not adequately considered when choosing the interest rate, meaning that such expectations are overshadowed by other mortgage-related variables. In other words, households appear to be erroneously more influenced by "external" parameters (beyond their control, such as the expected increase or decrease in rates) rather than by indicators intrinsic to their own financial situation, which could

also be partly influenced by the families their self. Also, we can consider that the sample could be affected by the bias highlighted by Zocchi in her analysis.

Risk attitude

Regarding the variable *Risk attitude*, in wave 2 the coefficients are not significant, whereas in wave 1 one of the coefficients is highly significant (falling within the 1% threshold) and another is weakly significant (within the 20% threshold). The highly significant coefficient refers to the category "take above average financial risks expecting to earn above average returns". It is strongly negative, implying that when a household considers itself more risk-taking than average, the probability of choosing a variable rate decreases compared to households defined as risk averse. In practical terms, an odds ratio of approximately 0.44 indicates that, holding all other conditions constant, the probability of opting for a variable rate decreases by 56% for households with a higher risk propensity relative to the reference category (risk averse). Moreover, the weakly significant coefficient associated with the category "moderate risk taker" is also negative, suggesting that even in this case the probability of choosing a variable rate decreases by 30% (with an odds ratio of approximately 0.70) compared to the fully risk-averse category.

Both results are contrary to expectations; in fact, it was hypothesized that as the propensity for risk increases the probability of choosing an ARM would also increase. It was expected that, relative to the reference category "risk averse", the coefficients would be positive and increasing for households with a higher risk propensity. These considerations are derived not only from the logic underlying the definition of the two types of interest rates but also from the evidence reported by Campbell and Cocco and by Coulibaly and Li, according to whom more risk-averse households tend to prefer a FRM.

A possible explanation for this contrary result is that households may not sufficiently consider their risk aversion when choosing the interest rate, instead prioritizing other objective parameters such as the trend in interest rates or the intrinsic characteristics of the mortgage (for example, the initial amount and the duration). The risk attitude, being more subjective and susceptible to recent events and personal disposition, might turn out to be less influential in the decision-making process.

In addition, the distortion may originate from a bias in the measurement of risk attitude. The questionnaire presented to households links risk attitude to return expectations, which could have influenced the responses. Some households might have indicated being less risk-averse than they actually were to avoid stating: "I am risk-averse, therefore I expect low returns from investments". As a result, some of them may have been mistakenly classified as above-average risk takers or moderate risk takers, despite actually being more cautious and more likely to choose a FRM.

Marital status

Regarding the variable *Marital status*, the coefficients are not significant in either sample. On one hand, considering a household in which the reference person is married, it might prefer an ARM as it is assumed that the combined income of the reference person and the spouse can sustain potential rate increases. On the other hand, the same household might choose a FRM in order to avoid compromising family expenses, which in many cases are neither reducible nor eliminable. Consequently, it is not entirely clear whether a positive or negative coefficient should be expected for *Marital status*.

On the theoretical front, the opinion of Campbell and Cocco is very explicit, as they argue that for a family with a married reference person the most obvious choice is an ARM, given a more stable or higher income compared to the one of a single or divorced person, thereby benefiting from both flexibility and potential accumulated savings. The lack of significance of the *Marital status* coefficient may result from the impact of this variable being overshadowed by the influence of other variables in the model.

Level of education

For the variable *Level of education*, the coefficients are not significant. It was hypothesized, for simplicity, that a higher level of education corresponds to greater financial knowledge, which one would expect to lead to a more informed interest rate choice. Therefore, a reference person with a high level of education (for example, a tertiary level) might, on one hand, prefer an ARM if they believe they can adequately manage risks or expect income growth; on the other hand, they might opt for a FRM if, thanks to a good understanding of the risks associated with an ARM and their own limitations, they decide to minimize potential uncertainties. It is therefore unclear whether one should expect a positive or negative coefficient.

In contrast, Ehrmann and Ziegelmeyer argue that for a more educated reference person the most obvious choice is an ARM, as income growth is expected and there is a greater ability to manage risks. The lack of significance of the coefficient associated with the level of education may therefore be due to its impact being overshadowed by the influence of other stronger variables in the model.

Occupational status

Regarding the variable *Occupational status*, in wave 1 the only (weakly) significant coefficient is the one relative to the "employee" category, while in wave 2 the weakly significant coefficient pertains to the "self-employed" category; in both cases the reference category is "other". The coefficients are both negative, indicating that when the reference person is an employee (in wave 1) or self-employed (in wave 2), the household tends to choose a variable rate less frequently compared to the reference category. In wave 1, an odds ratio of about 0.65 implies that the probability of choosing a variable rate decreases by 35% for a reference person classified as employee; similarly, in wave 2, an odds ratio of about 0.55 suggests a 45% decrease in the probability of opting for a variable rate if the reference person is self-employed.

Since the coefficients related to *Occupational status* are not consistent between the two waves, it is complex to determine which category between "employee" and "self-employed" is actually more inclined to choose a fixed or variable rate. Theoretically, according to Campbell and Cocco, a self-employed worker should opt for an ARM, as it allows to benefit from a lower initial cost compared to a FRM. However, it is not guaranteed that such a choice is always advantageous: a self-employed worker might, in fact, lack the income stability necessary to cope with potential increases in instalments, leading them to prefer a FRM. Regarding the employee, although possessing a generally stable and non-growing income, they might still opt for an ARM if their accumulated savings or their income level allows them to manage interest rate fluctuations. These scenarios highlight that the relationship between occupational status and the choice of interest rate is complex and not unequivocal. Finally, the low significance of the coefficients in the model may be due to the impact of occupational status being overshadowed by the influence of other, more relevant variables, making it difficult to isolate its effect.

Standardized parameters

In this section, only the impacts of the variables with significant coefficients (up to a maximum threshold of 20%) are discussed. The effect of each variable on the outcome varies considerably between the two waves; therefore, these impacts are examined separately before arriving at a shared conclusion.

Sourco	Value	Standard	Dr. > Chi2
300100	value	error	PIZCIII
Initial amount	-0,214	0,087	0,014
Duration	-0,037	0,071	0,606
Total real assets	0,005	0,070	0,946
Age	-0,020	0,075	0,790
Purpose-1	0,007	0,068	0,914
Purpose-2	0,000	0,000	
Year of signing-1	-0,191	0,071	0,007
Year of signing-2	-0,247	0,073	0,001
Year of signing-3	0,077	0,071	0,280
Year of signing-4	0,000	0,000	
Expenses compared to income-1	-0,048	0,068	0,484
Expenses compared to income-2	-0,028	0,067	0,679
Expenses compared to income-3	0,000	0,000	
Future income expectations-1	-0,080	0,068	0,241
Future income expectations-2	0,025	0,065	0,697
Future income expectations-3	0,000	0,000	
Risk attitude-1	-0,009	0,061	0,883
Risk attitude-2	-0,170	0,065	0,008
Risk attitude-3	-0,095	0,065	0,140
Risk attitude-4	0,000	0,000	
Marital status-1	-0,004	0,061	0,952
Marital status-2	0,000	0,000	
Level of education-1	0,066	0,084	0,429
Level of education-2	0,001	0,078	0,994
Level of education-3	0,000	0,000	
Occupational status-1	-0,114	0,083	0,172
Occupational status-2	-0,094	0,077	0,225
Occupational status-3	0,000	0,000	

Fig.3 - Standardized parameters wave 1 (Variable interest rate as the reference category)

For wave 1, the variables with the greatest impact are the reference years and the initial amount borrowed, while those with the least impact are risk attitude and occupational status. It is interesting to note that the variable *Initial amount*, which appeared almost negligible in the model coefficients, turns out to be among the most impactful when standardized. This highlights how its relative importance was partly obscured by the unit of measurement, and how the standardized coefficient offers a clearer view of its impact on the mortgage choice. Moreover, for the more impactful variables the results are substantially as expected, whereas for *Risk attitude* the outcomes are contrary to the predictions.

Source	Value	Standard error	Pr > Chi ²
Initial amount	-0,116	0,086	0,176
Duration	-0,168	0,070	0,016
Total real assets	0,151	0,076	0,047
Age	-0,056	0,070	0,421
Purpose-1	0,101	0,063	0,110
Purpose-2	0,000	0,000	
Year of signing-1	-0,010	0,064	0,875
Year of signing-2	-0,135	0,070	0,052
Year of signing-3	0,045	0,068	0,509
Year of signing-4	0,000	0,000	
Expenses compared to income-1	0,051	0,061	0,398
Expenses compared to income-2	0,074	0,062	0,232
Expenses compared to income-3	0,000	0,000	
Future income expectations-1	-0,174	0,126	0,168
Future income expectations-2	-0,088	0,127	0,490
Future income expectations-3	0,000	0,000	
Risk attitude-1	-0,078	0,077	0,314
Risk attitude-2	-0,067	0,063	0,290
Risk attitude-3	-0,005	0,061	0,930
Risk attitude-4	0,000	0,000	
Marital status-1	0,033	0,059	0,581
Marital status-2	0,000	0,000	
Level of education-1	0,006	0,082	0,945
Level of education-2	-0,071	0,076	0,352
Level of education-3	0,000	0,000	
Occupational status-1	-0,023	0,083	0,780
Occupational status-2	-0,118	0,079	0,136
Occupational status-3	0,000	0,000	

Fig.4 - Standardized parameters wave 2 (Variable interest rate as the reference category)

For wave 2, the variables with the greatest impact on the outcome are *Future income expectations*, *Duration*, and *Total real assets*; the *Year of signing* falls in the middle of the ranking, resulting in an average impact. Among the less impactful variables there are *Occupational status*, *Initial amount*, and *Purpose*. Again, two of the variables that exhibited low model coefficients (*Duration* and *Total real assets*) subsequently turned out to be particularly impactful; for two of the three most impactful variables the results are substantially as expected, while the variable *Purpose* shows a result contrary to what was anticipated.

Considering the impact rankings in both waves, it can be deduced that the *Year of signing* is particularly determinative. In contrast, *Occupational status*, that represents a characteristic of the reference person, has a limited impact on the outcome.

Model performance

Confusion matrix



The confusion matrix is a table that displays the comparison between the values predicted by the model and the actual values of the target variable, and it is used to evaluate the performance of a logistic regression model with a binary dependent variable. In this table, each quadrant shows the number of observations correctly or incorrectly classified based on the chosen mortgage rate. In the top left quadrant, "True Positives" (TP) are displayed, representing the number of mortgages for which the model correctly predicted a variable rate. The bottom right quadrant shows the "True Negatives" (TN), corresponding to the number of mortgages accurately predicted as FRM. In the top right quadrant, "False Negatives" (FN) are indicated, referring to the mortgages for which the model wrongly predicted a fixed rate. Finally, the bottom left quadrant represents the "False Positives" (FP), which are the mortgages for which the model erroneously predicted an ARM. Next to the TP and TN, the sum for each row is indicated. It should be noted that the total number of observations considered in the regression is 407 for wave 1 and 419 for wave 2; these numbers are slightly lower than those in the chosen subset, because XLSTAT requires all variables to be present in order to include an observation.

It is interesting to note the difference between the two matrices: in wave 1 the FN and FP are almost equal, whereas in wave 2 the FP are almost twice the FN, which indicates a tendency of the model to predict a greater number of variable rates than actually observed. This imbalance may partly derive from the data distribution, since the percentage of ARMs is 55% in wave 2 compared to 50% in wave 1, thereby influencing the model toward a higher number of FP. However, the fact that the FP are nearly twice the FN suggests that the model exhibits an additional bias in favour of the variable rate class.

Some summary metrics are reported below:

• Accuracy

Calculated by the formula: (TP + TN) / Total observations. It is 0.639 in wave 1 and 0.63 in wave 2.

• Sensitivity (or TP rate)

Calculated by the formula: TP / (TP + FN). It is 0.639 in wave 1 and 0.762 in wave 2.

• Precision

Calculated by the formula: TP / (TP + FP). It is 0.635 in wave 1 and 0.644 in wave 2.

These results indicate that, in wave 1, the model correctly classifies a slightly higher percentage of observations compared to wave 2, while in wave 2 the model is more sensitive, that is, more effective in recognizing the True Positives, while not being less precise.

ROC Curve



The ROC Curve is a graph that illustrates the trade-off between "sensitivity" (TP rate) and "1 - specificity" (FP rate) of the model at different classification thresholds. It is an additional tool for evaluating the performance of a logistic regression model with a binary dependent variable. A perfect model is characterized by a curve that rises rapidly along the vertical axis (indicating high sensitivity) and shifts only slightly horizontally (indicating high specificity). The AUC (Area Under the Curve) provides a summary of the overall quality of the model: a value of 0.5 indicates random prediction, while values close to 1 indicate a good ability to distinguish between the two classes (variable and fixed). When the curves from the two waves are overlaid, they intersect, suggesting that at certain thresholds the model performs better in wave 1, while at others it performs better in wave 2. However, when considering the evaluation across all possible thresholds, the higher AUC in wave 1 leads to the conclusion that the model applied to that sample performs better, that is, it is more effective at distinguishing between ARMs and FRMs.

Chapter 4 - Conclusions

4.1 Main results

Among the results obtained from the analysis, only those that meet at least the 5% significance threshold are reported.

First, as the initial amount increases, the probability of choosing a variable rate slightly decreases. This result is consistent with several texts in the reference literature and follows common logic, considering that managing rate fluctuations when the payment is larger can be more difficult. Furthermore, it demonstrates that the bias highlighted by the author Paola Zocchi, according to which families erroneously prefer an ARM as the amount increases, is not observed in the sample.

Second, as the mortgage duration increases, the probability of choosing a variable rate slightly decreases. Although this result contrasts with what is reported in some texts in the literature, it is coherent when considering the positive correlation observed between the initial amount and the mortgage duration.

Third, as the real value of the family's assets increases, the probability of opting for a variable rate slightly increases. This result is in line with the analysed literature if real assets are considered as a representation of the family's economic stability; however, since real assets provide financial security but are not easily convertible into liquidity, this variable should be considered together with the family's income in order to have a clear picture of the capacity to bear potential rate increases.

Another result indicates that when the mortgage was taken out between the beginning of 2001 and the end of 2005, it is more likely to be a FRM, whereas when the mortgage was signed in the post-global financial crisis period (between 2008 and 2010) it is more likely to be an ARM. These results suggest that the uncertainty of the historical period influences the choice but does not necessarily push towards the safer option represented by a FRM, since what matters most is the forecast of future rate trends.

Finally, an unexpected result concerns the analysis of the variable *Risk attitude*. It was supposed that a lower risk aversion would lead to a higher probability of choosing an ARM. The fact that this does not occur could be explained mainly by two considerations. On one side, families do not adequately consider this aspect when choosing between a

fixed and a variable rate, instead favouring other factors. This is confirmed by the standardized parameters from which we understand that the relative impact of risk attitude on the outcome is not much significant. On the other side, there could be a bias originating from the association between risk attitude and return expectations in the questionnaire, leading to an erroneous classification of families in risk-loving categories.

It should be noted also that this parameter describes the family at the time of the survey and not at the time of the mortgage signing; therefore, it is not possible to determine whether the response to the question would have been the opposite at the time of the choice.

4.2 Limitations

Some limitations encountered were unavoidable, while the assumptions adopted were essential, either because they were imposed or needed to simplify the analysis.

Among these, first, is the fact that the analysis was limited to the choice between fixed and variable rates, without considering intermediate mortgage types or renegotiation instruments.

Second, the sample was based on a data pool from a single geographic area (Italy) and on a restricted time period. In order to compare the two waves, only the common years had to be selected, and the sample was further reduced to avoid responses related to mortgages opened more than 13 years before the survey. As mentioned in the conclusions, some responses are inevitably correlated with the family's situation at the time of the survey; to treat all variables as true predictors of the outcome, it would have been necessary to consider exclusively data relating to the moment of the choice. This is a condition met for some variables: *Initial amount, Duration, Age, Purpose*, and *Year of signing*. While for others it was not possible, rendering these variables less reliable when changes occurred between the two moments.

Third, behavioural biases in households and needs that do not emerge in the model were identified. For instance, a household might not be able to account for all the variables that characterise it when choosing the interest rate type, or it might be influenced by non-objective parameters. Furthermore, the choice might be conditioned by the bank's opinion or strategy, which, on one hand, will tend to act in the client's best interest, while on the other hand, it will avoid incurring in losses.

Finally, similar to the analysed articles, this analysis assumes that the decision regarding the interest rate type is made sequentially, which means, it is made only after determining other parameters, such as the initial amount and the duration.

4.3 Further studies

Among the improvements that could be incorporated in future studies, the first is the need to increase the sample considered. Although the initial sample was rather large, the chosen subset obtained is reduced. It would be particularly useful to employ data concentrated within a specific period, in order to better relate the interest rate choice to the economic conditions at that moment, particularly regarding the difference between variable and fixed rates. Moreover, it would be interesting to repeat the analysis in a more recent period.

Another improvement could consist in increasing the number of variables considered to obtain a more complete model. In particular, some income-related variables could have provided a more precise indication of the family's ability to manage potential rate increases. These variables are as follows:

- 1. *Total household gross income*, that is, the sum of all possible sources of income, such as employment, any private business, investments, rents from properties, etc.
- 2. *Debt-to-income ratio*, which represents the ratio between total liabilities and total household gross income, providing a clearer indication of the financial stress to which the family is subjected, as very high income may correspond to significant liabilities.
- 3. *Stability (or risk) of income*, which assesses the income sources and their variability over time, offering an indication not only on the amount of money or assets available, but also on their solidity or volatility.
- 4. *Relative size of mortgage*, meaning the mortgage amount relative to the income at the time of signing, which allows for a clearer determination of whether an initial amount is high or modest, depending on the family's capacity to sustain a certain expenditure.

Finally, in addition to the income-related variables, it would be advisable to consider certain aspects of behavioural finance, evaluating how psychological traits and biases might influence the choice of interest rate.

4.4 Final considerations

In conclusion, despite the limitations encountered during the analysis, several interesting observations have been reached. In some cases, they even contradict the expectations derived from the literature. In particular, the bias proposed by the author Zocchi, suggesting that an increase in the mortgage amount would push families to choose an ARM, was not observed, nor is there evidence that a longer mortgage duration leads to a more frequent choice of an ARM. Furthermore, it emerged that historical uncertainty does not necessarily direct the decision toward a FRM; rather, the decision appears to be more influenced by forecasts of future rate trends. Finally, the most counterintuitive result concerns risk attitude: it appears that as risk aversion increases, the probability that a family chooses a variable rate also increases.

Appendix A

A.1 Descriptive statistics at person level - Quantitative variables

Full sample

Age (full sample)	W1	W2
Nbr. of observations	19836	19366
Minimum	0	0
Maximum	85	85
Median	47	51
Mean	45	48
Standard deviation (n-1)	23	23

Age (chosen subset)	W1	W2
Nbr. of observations	425	439
Minimum	26	26
Maximum	83	85
Median	45	48
Mean	47	49
Standard deviation (n-1)	10	10

A.2 Descriptive statistics at person level - Qualitative variables

Full sample

Statistic at person level (full sample)	Nbr. Obs. W1	Nbr. Obs. W2	Categories	Rel. freq. per category W1 (%)	Rel. freq. per category W2 (%)
Gender	19836	19366	Female	52	52
			Male	48	48
Marital status	17274 ن	17171	Single	27	28
			Married	59	56
			Divorced	4	5
			Widowed	10	11
Level of education	17274	17171	Primary or below	23	22
			Lower secondary	31	31
			Upper secondary	35	35
			Tertiary	11	12
Occupational status	17274	17171	Employee with a permanent position	27	24
			Employee with a temporary contract	5	5
			Self-employed (without employees)	6	6
			Self-employed (with employees)	2	2
			Retiree or Early retiree	31	27
			Student or Unpaid intern	8	8
			Unemployed	8	10
			Perform domestic tasks	13	11
			Permanently disabled	0	2
			Unpaid family worker	0	C
			Other	0	5

Statistic at person level (chosen subset)	Nbr. Obs. W1	Nbr. Obs. W2	Categories	Rel. freq. per category W1 (%)	Rel. freq. per category W2 (%)
Gender	425	439	Female	37	42
			Male	63	58
Marital status	425	439	Single	9	12
			Married	82	76
			Divorced	7	9
			Widowed	2	3
Level of education	425	439	Primary or below	7	5
			Lower secondary	27	28
			Upper secondary	43	45
			Tertiary	23	22
Occupational status	425	439	Employee with a permanent position	59	57
			Employee with a temporary contract	3	4
			Self-employed (without employees)	11	13
			Self-employed (with employees)	4	2
			Retiree or Early retiree	13	12
			Student or Unpaid intern	0	0
			Unemployed	3	5
			Perform domestic tasks	7	6
			Permanently disabled	0	1
			Unpaid family worker	0	C
			Other	0	0

A.3 Descriptive statistics at family level - Quantitative variables

Full sample

Statistic (W1 full sample)	Current value all properties	Total real assets	Outstanding balance of household's liabilities	Initial amount	Duration
Nbr. of observations	7.951	7.951	7.951	626	619
Minimum	-	-	-	1.900	2
Maximum	5.700.000	26.005.000	1.320.000	1.000.000	40
Median	156.000	181.000	-	80.000	18
Mean	223.092	264.448	9.970	102.311	18
Standard deviation (n-1)	330.168	521.252	42.701	90.347	7

Statistic (W2 full sample)	Current value all properties	Total real assets	Outstanding balance of household's liabilities	Initial amount	Duration
Nbr. of observations	8.156	8.156	8.156	644	642
Minimum	-	-	-	4.000	2
Maximum	3.628.000	13.600.000	872.000	720.000	40
Median	150.000	160.000	-	100.000	20
Mean	195.384	227.003	7.427	110.720	21
Standard deviation (n-1)	254.308	347.189	32.478	68.684	7

Statistic (W1 chosen subset)	Current value all properties	Total real assets	Outstanding balance of household's liabilities	Initial amount	Duration
Nbr. of observations	425	425	425	425	425
Minimum	30.000	48.000	1.200	2.500	3
Maximum	2.626.100	3.025.000	795.000	900.000	40
Median	240.000	260.600	76.300	100.000	20
Mean	333.126	374.404	94.014	110.427	19
Standard deviation (n-1)	313.667	387.725	80.806	78.207	7

Statistic (W2 chosen subset)	Current value all properties	Total real assets	Outstanding balance of household's liabilities	Initial amount	Duration
Nbr. of observations	439	439	439	439	439
Minimum	20.000	21.500		15.000	5
Maximum	1.675.000	3.343.500	739.000	720.000	40
Median	200.000	210.000	60.000	100.000	20
Mean	255.756	285.318	70.982	110.423	22
Standard deviation (n-1)	204.759	266.644	68.856	65.106	7

A.4 Descriptive statistics at family level - Qualitative variables

Full sample

Statistic at family level (full sample)	Nbr. Obs. W1	Nbr. Obs. W2	Categories	Rel. freq. per category W1 (%)	Rel. freq. per category W2 (%)
HMR tenure status	7951	8156	Own all	69	70
			Own part	2	2
			Rented / Sublet	20	19
			Free use	9	g
Purpose	626	644	To purchase the HMR	72	79
			Other	28	21
Year of signing	605	644	Before 2001	20	10
			2001-2002	15	8
			2003-2005	25	2:
			2006-2007	15	18
			2008-2010	25	23
			After 2010	0	20
Expenses compared to income (last 12 months)	7951	8156	Expenses about the same as income	56	27
			Expenses higher than income	8	9
			Expenses lower than income	36	64
Mortgages using HMR as collateral	5636	5875	No	89	89
			Yes	11	11
Mortgages using other properties as collateral	1933	1881	No	94	91
			Yes	6	
Future income expectations	7257	7644	More than prices	12	4
			Less than prices	57	54
			About the same as prices	31	
Risk attitude	7951	8156	Take substantial financial risks expecting to earn substantial returns	1	1
			Take above average financial risks expecting to earn above average returns	18	11
			Take average financial risks expecting to earn average returns	32	29
			Not willing to take any financial risk	49	55

Statistic at family level (chosen subset)	Nbr. Obs. W1	Nbr. Obs. W2	Categories	Rel. freq. per category W1 (%)	Rel. freq. per category W2 (%)
HMR tenure status	425	439	Own all	98	99
			Own part	2	1
			Rented / Sublet	0	0
			Free use	0	0
Purpose	425	439	To purchase the HMR	78	83
			Other	22	17
Year of signing	425	439	Before 2001	0	C
			2001-2002	19	12
			2003-2005	31	31
			2006-2007	19	26
			2008-2010	31	31
			After 2010	0	C
Expenses compared to income (last 12 months)	425	439	Expenses about the same as income	53	26
			Expenses higher than income	14	14
			Expenses lower than income	33	60
Future income expectations	407	419	More than prices	9	34
			Less than prices	64	60
			About the same as prices	27	6
Risk attitude	425	439	Take substantial financial risks expecting to earn substantial returns	2	1
			Take above average financial risks expecting to earn above average returns	17	11
			Take average financial risks expecting to earn average returns	37	33
			Not willing to take any financial risk	44	55

Appendix B

Correlation matrix

2																															
Correlation matrix (W I):																															
	Initial		Current	Total real	Outstand in g bal. of				Year of	Year of	Year of	Year of	Expenses I	Expenses	Expenses	Future	Future	Future	Risk	Risk	Risk	Risk	Martal	Marital	Level of	Level of	Level of 0	ccupation O	coupation Ox	ccup ation	Interest
	amount		properties	assets	household" s liab.	ą	a about a	- appear	signing-1	sign ing-2	signing-3	signing-4 to	o income-1 to	income-2 to	income-3	s-1	s-2	s-3	attitude-1 a	ttitud e-2	attitude-3 a	attitude-4	status-1	status-2 e	ducation-1 e	ducation-2 e	ducation-3 a	i status-1 a	l status-2 a	91 status-3	rate
n itial a mo unt	1	0,409	0,296	0,299	0,771	-0,065	0,205	-0,205	-0,145	-0,012	0,119	0,087	0,020	-0,080	0,071	0,139	-0,080	-0,003	0,059	0,007	0,001	-0,022	0,057	-0,057	-0,117	-0,086	0,230	-0,020	0,178	-0,131	-0,159
Duration		-	-0,096	-0,067	0,420	-0,263	0,348	-0,348	-0,074	-0,001	0,131	-0,046	0,030	-0,049	0,030	0,077	-0,021	-0,027	-0,029	0,028	-0,039	0,025	-0,043	0,043	-0,095	0,054	0,042	0,163	-0,005	-0,184	-0,094
Current value all properties			-	0,938	0,303	0,173	-0,196	0,196	-0,064	-0,072	0,047	0,089	0,042	-0,065	0,038	0,035	-0,056	0,037	0,060	-0,080	0,060	-0,014	0,074	-0,074	-0,180	-0,117	0,335	-0,130	0,161	0,011	-0,014
fotal real assets				-	0,429	0,133	-0,186	0,186	-0,070	-0,092	0,100	0,068	0,072	-0,089	0,042	0,057	-0,067	0,035	0,056	-0,059	0,039	-0,008	0,057	-0,057	-0,198	-0,091	0,324	-0,170	0,245	-0,014	-0,025
Outstanding bal of household's liab.					-	-0,075	0,162	-0,162	-0,219	-0,107	0,185	0,142	0,006	-0,056	0,056	0,117	-0,072	0,002	0,049	0,017	-0,052	0,024	0,005	-0,005	-0,153	-0,064	0,244	-0,029	0,225	-0,161	-0,101
ke						-	-0,359	0,359	-0,164	-0,043	-0,076	0,251	0,077	0,007	-0,064	0,001	0,048	-0,053	-0,013	-0,042	-0,039	0,074	-0,078	0,078	0,206	-0,140	-0,063	-0,419	0,037	0,452	0,082
ourpose-1							-	-1,000	0,093	0,060	0,018	-0,156	-0,061	6000	0,036	-0,033	0,043	-0,025	800,0-	0,001	0,023	-0,021	0,001	-0,001	-0,080	0,008	0,079	0,148	-0,057	-0,122	-0,067
ruipose-z								-	-0,090	0000-	010,0-	001,0	100,0	- ,	-0,030	0051	0,040	0,013	0000	-0,001	C20(0-	120/0	0,001	0,001	0,000	-0,000	-0,01	-0,140	-0.037	2002	1000
rear of signing-2										-	-0, 32 3	-0,443	-0,096	-0,02.8	0,100	0,045	0,068	-0,103	0,019	-0,072	-0,020	0,069	0,006	-0,006	0,001	0,080	-0,037	0,014	-0,087	0,059	-0,148
rear of signing-3												-0,315	0,106	-0,02.6	-0,049	0,024	-0,057	0,046	-0,022	0,012	0,065	-0,067	0,083	-0,033	0,048	-0,071	0,031	-0,084	0,105	0,007	0,135
rear of signing-4												•	-0,025	0,059	-0,044	-0,022	-0,036	0,054	-0,055	0,078	-0,015	-0,030	-0,046	0,046	-0,041	0,006	0,038	0,030	0,010	-0,043	0,097
expenses compared to income-1													-	-0,429	-0,274	0,025	0,015	-0,032	-0,004	-0,068	0,026	0,028	0,003	-0,003	0,035	-0,048	0,017	-0,113	-0,013	0,141	0,004
expenses compared to income-2															-0,751	-0,137	0,092	-0,011	0,024	0,034	-0,062	0,028	0,083	-0,083	0,171	-0,072	-0,103	-0,026	0,033	0,002	0,031
Expenses compared to income-3															-	0,128	-0,109	0,035	-0,022	0,013	0,047	-0,050	-0,090	0,090	-0,208	0,112	0,097	0,110	-0,025	-0,106	-0,037
uture income expectations-1																-	-0,418	-0,194	0,201	-0,033	-0,029	-0,003	-0,004	0,004	-0,085	0,028	0,005	-0,029	0,122	-0,072	-0,096
uture income expectations-2																		-0,810	-0,114	-0,029	-0,090	0,141	-0,043	0,043	0,051	-0,078	0,035	0,037	-0,108	0,051	0,057
uture income expectations-3																		-	-0,007	0,053	0,115	-0,151	0,049	-0,049	-0,082	0,066	-0,041	-0,022	0,039	-0,008	0,001
vskatutuoe-1																			-	-0,05	-0,108	-0,125	0,067	-0,067	-0,060	-0,019	0,088	-0,032	0,085	-0,036	-0,036
tisk attitude-3																				•	1	-0,674	-0,110	0,110	-0,047	0,058	-0,017	-0,018	0,062	-0,032	-0,016
lisk attitude-4																						-	0,074	-0,074	0,050	-0,097	0,059	-0,070	-0,054	0,128	0,102
Marital status-1																							-	-1,000	0,056	0,002	-0,064	0,051	-0,108	0,035	-0,003
Marital status-2																									-0,056	-0,002	0,064	-0,051	0,108	-0,035	0,003
evel of education-1																									1	-0,613	-0,383	-0,153	-0,124	0,284	0,095
.evel of education-2																											-0,495	0,131	-0,018	-0,136	-0,056
evel of education-3																												0,015	0,157	-0,153	-0,039
Occupational status-1																												-	-0,542	-0,688	-0,075
Occupational status-2																													-	-0,237	-0,044
Occupational status-3																														-	0,124
n terest rate																															-

Correlation matrix (W2):																															
	Initial		Current	Total real	Outstand in a bal. of				Year of	Year of	Year of	Year of	Expenses	Expenses E	Expenses	Future	Future	Future	Risk	Risk	Risk	Risk	Martal	Marital	Level of	Level of	Level of 0	ccupation 0	vcupation 0	ecup ation	Interest
	amount	Dui atron	properties	assets	household" s liab.	Ą	raiboser	ruipose-2	igning-1	sign ing-2	signing-3	signing-4 to	b income-1 to	b income-2 to	income-3	s-1	pectation e: s-2	s-3	nttitude-1 a	ottitude-2 a	ottitude-3	ntitude-4	status-1	status-2 e	ducation-1 e	fucation-2 e	ducation-3 a	ol status-1 a	Istatus-2 a	il status-3	rate
In itial amo unt	-	0,422	0,453	0,514	0,633	-0,076	0,200	-0,200	-0,137	-0,113	0,177	0,087	0,010	-0,011	0,003	0,036	-0,015	-0,044	0,113	0,010	0,049	-0,079	0,096	-0,096	-0,132	-0,038	0,192	-0,029	0,244	-0,172	-0,080
Duration		-	0,002	0,042	0,491	-0,242	0,239	-0,239	-0,040	-0,074	0,086	0,020	-0,021	0,039	-0,020	0,065	-0,050	-0,030	0,103	-0,008	-0,029	0,009	-0,029	0,029	0,015	-0,019	0,006	0,143	-0,018	-0,149	-0,141
Current value all properties			-	0,922	0,383	0,168	-0,169	0,169	-0,007	-0,054	0,019	0,040	0,072	0,161	-0,196	0,050	-0,066	0,038	0,092	0,087	0,140	-0,207	0,036	-0,036	-0,159	-0,052	0,238	-0,129	0,189	-0,010	0,029
Total real assets				1	0,515	0,129	-0,126	0,126	-0,017	-0,070	0,021	0,061	0,049	0,142	-0,163	0,073	-0,081	0,023	0,083	0,116	0,127	-0,211	0,049	-0,049	-0,151	-0,021	0,192	-0,161	0,258	-0,031	0,036
Outstanding bal of household's liab.					1	-0,069	0,070	-0,070	-0,185	-0,188	680'0	0,228	0,007	0,014	-0,017	0,057	-0,035	-0,044	0,064	0,049	0,078	-0,120	-0,007	0,007	-0,093	-0,023	0,131	-0,040	0,205	-0,126	-0,044
Age						-	-0,225	0,225	-0,104	-0,036	-0,028	0,133	0,044	-0,085	0,046	0,008	-0,009	0,002	-0,032	0,02.8	0,020	-0,029	-0,055	0,055	0,206	-0,079	-0,134	-0,397	-0,059	0,506	0,009
Purpose-1							1	-1,000	0,062	0,073	-0,018	-0,097	-0,004	-0,089	0,038	0,050	-0,035	-0,029	0,002	0,026	-0,050	0,032	0,029	-0,029	-0,036	-0,059	0,111	0,098	0,029	-0,137	0,008
Purpose-2 Vant of similar 1								-	-0,062	-0,073	810/0	7.60'0	0,004	0,039	0,0-8	-0,050	0,035	0,029	-0,002	-0,026	0,050	-0,032	-0,029	0,029	0,036	0,059	-0,111	-0,098	-0,029	0,137	-0,008
Year of signing-2										-	-0, 398	-0,457	0,020	-0,044	0,026	-0,062	0,093	-0,071	0,049	-0,018	-0,063	0,058	0,054	-0,054	0,041	-0,012	-0,030	-0,004	-0,076	0,068	-0,094
Year of signing-3											-	-0,404	-0,045	-0,063	0,088	0,075	-0,062	-0,023	0,066	-0,017	0,015	-0,020	0,117	-0,117	-0,068	0,017	0,056	-0,027	0,090	-0,045	0,034
Year of signing-4												•	-0,016	0,086	-0,066	0,025	-0,052	0,061	-0,082	0,046	0,013	-0,021	-0,174	0,174	0,030	-0,027	-0,002	0,019	0,005	-0,026	0,027
Expenses compared to income-1														-0,240	-0,483	-0,097	0,081	0,027	0,128	-0,040	-0,076	0,066	-0,012	0,012	0,095	-0,075	-0,015	-0,161	-0,009	0,192	0,042
Expenses compared to income-2														-	-0,735	0,051	-0,103	0,115	-0,027	0,01.4	0,118	-0,114	-0,136	0,136	-0,114	-0,053	0,188	0,031	0,037	-0,066	0,063
Expenses compared to income-3															-	0,021	0,036	-0,122	-0,065	0,015	-0,053	0,056	0,131	-0,131	0,036	0,100	-0,159	0,085	-0,027	-0,074	-0,086
Future income expectations-1																-	-0,889	-0,174	0,082	0,093	0,006	-0,081	-0,011	0,011	-0,040	-0,067	0,124	-0,021	0,095	-0,056	-0,083
Future income expectations-2																	-	-0,296	-0,107	-0,102	0,012	0,075	-0,021	0,021	0,061	0,052	-0,130	0,002	-0,114	0,093	0,052
Future income expectations-3																		-	0,059	0,02.4	-0,039	0,008	0,067	-0,067	-0,048	0,02.8	0,020	0,038	0,047	-0,084	0,061
Risk attitude-1																			-	-0,040	-0,086	-0,133	0,070	-0,070	-0,081	0,049	0,031	0,012	0,063	-0,067	-0,066
Risk attitude-2																				-	-0,238	-0,368	0,065	-0,065	-0, 155	0,089	0,066	-0,048	0,107	-0,034	-0,055
Risk attitude-3																					-	-0,786	0,016	-0,016	-0,041	-0,043	0,097	0,134	-0,098	-0,071	0,042
Risk attitude-4																						-	-0,070	0,070	0,152	-0,024	-0,139	-0,100	0,013	0,104	0,010
Marital status-1																							-	-1,000	0,066	-0,037	-0,029	0,033	-0,023	-0,019	0,023
Marital status-2																								-	-0,066	0,037	0,029	-0,033	0,023	0,019	-0,023
Level of education-1																									-	-0,623	-0,363	-0,216	-0,105	0,336	0,041
Level of education-2																										-	-0,503	0,099	0,017	-0,129	-0,069
Level of education-3																											-	0,121	0,095	-0,218	0,038
Occupational status-1																												-	-0,530	-0,703	0,012
Occupational status-2																													-	-0,230	-0,071
Occupational status-3																														-	0,045
In terest rate														_												_					-

Bibliography

- Banca d'Italia, Indagine sui bilanci delle famiglie italiane.
- https://www.bancaditalia.it/statistiche/tematiche/indagini-famiglieimprese/bilanci-famiglie/dati-indagine-europea/index.html
- https://www.bancaditalia.it/pubblicazioni/metodi-e-fonti-note/metodi-note-2022/MOP_IBF_ita.pdf
- https://www.bancaditalia.it/statistiche/tematiche/indagini-famiglieimprese/bilanci-famiglie/faq-famiglie/index.html
- "Household Risk Management and Optimal Mortgage Choice" by John Y. Campbell and João F. Cocco, 2003.
- "Why do Italian households prefer adjustable rate mortgages?" by Paola Zocchi, 2013.
- "Mortgage Choice in the Euro Area: Macroeconomic Determinants and the Effect of Monetary Policy on Debt Burdens" by Michael Ehrmann and Michael Ziegelmeyer, 2017.
- "Choice of Mortgage Contracts: Evidence from the Survey of Consumer Finances" by Brahima Coulibaly and Geng Li, 2007.
- https://www.mutuionline.it/guide-mutui/introduzione-ai-mutui.asp
- https://www.mutui.it/guida-mutui/tasso-irs-mutui-come-funziona.html
- https://www.euribor-rates.eu/it/tasso-interesse-bce/
- https://grafici.altervista.org/inflazione-tassi-di-interesse-e-di-cambio-in-italia/
- https://www.istat.it/60annidieuropa/lavoro.html
- https://www.diritto.it/il-nuovo-concetto-legale-di-famiglia-dopo-l-approvazionedella-legge-n-76-del-20-maggio-2016/
- https://www.treccani.it/enciclopedia/la-crisi-finanziaria-e-l-italia_(Il-Librodell'Anno)/
- https://www.bancaditalia.it/pubblicazioni/indagine-inflazione/2006-indagine-inflazione/06/index.html
- https://www.rgs.mef.gov.it/_Documenti/VERSIONE-I/Attivit- i/Contabilit_e_finanza_pubblica/Archivio-d/Relazione-1/2007/1.2.-L-Economia italiana-nel-2006.pdf
- https://www.questionpro.com/blog/it/matrice-di-correlazione/
- https://riani.unipr.it/sde/SDE03.pdf

- Regressione Multipla e Regressione Logistica: concetti introduttivi ed esempi I Edizione – ottobre 2016 – Vincenzo Paolo Senese
- https://www.ibm.com/it-it/topics/logistic-regression
- https://adrianogilardone.com/regressione-logistica/
- https://psiclab.altervista.org/MetTecPsicClinica2015/Materiali/_3_Regr_Logistic
 a_2014_2015.pdf
- https://stem.elearning.unipd.it/pluginfile.php/1018432/mod_resource/content/2/ Lezione_Logistic_finale.pdf
- https://www.statology.org/interpret-logistic-regression-intercept/
- https://my.liuc.it/MatSup/2018/A86051/Lezione_10_AA201819.pdf
- https://my.liuc.it/MatSup/2016/A86051/esercitazione_11_aa1617.pdf
- https://www.ibm.com/it-it/topics/confusion-matrix
- https://www.andreaprovino.it/the-roc-curve-receiver-operating-characteristic