

Household Finance in Greece

Freideriki Plesiti

A dissertation submitted in partial fulfillment of the requirements for the degree of Master of Science in Applied Economics & Data Analysis

School of Economics and Business Administration

Department of Economics

Master of Science

"Applied Economics and Data Analysis"

September 2023

University of Patras, Department of Economics

Freideriki Plesiti

© 2023 − All rights reserved

Dissertation Committee

Research Supervisor:	Eleftherios Goulas	Assistant Professor
Dissertation Committee Member:	Nicholas Giannakopoulos	Assistant Professor
Dissertation Committee Member:	Georgios Filis	Assistant Professor

The present dissertation entitled

"Household Finance in Greece"

was submitted by **Freideriki Plesiti, ID 1104212** , in partial fulfillment of the requirements for the degree of Master of Science in *"Applied Economics & Data Analysis"* at the University of Patras and was approved by the Dissertation Committee Members.

I would like to dedicate my dissertation to my dearest family and friends. Thank you for the unwavering encouragement and support.

Acknowledgments

I would like to express my deepest gratitude to my supervisor, Eleftherios Goulas. His patient guidance, motivation and advice have been instrumental in my academic journey this semester. I extend my gratitude to the members of my dissertation committee, Nicholas Giannakopoulos and Georgios Filis, for honoring me with their presence on my dissertation committee.

Abstract

Using micro data from the Greek Survey on Statistics of Income and Living Conditions (EU-SILC), this dissertation attempt to identify the constraints households during the period 2007-2022, emphasizing on the effects of several factors on the probability of default on mortgage or rent repayments. The study not only acknowledges the importance of default risk but also recognizes the broader spectrum of financial challenges that households might encounter, particularly in the realm of housing expenses. Based on the results of the econometric estimates and considering various demographic characteristics, household characteristics as well as regional effects, it has been found that lower default levels are observed for high-income households. As a consequence, this study had certain value for the prevention and control of constraint households.

Keywords: Greece, Households, mortgage, rent repayment, probability default

Περίληψη

Χρησιμοποιώντας δεδομένα σε μικροεπίπεδο, από την Έρευνα Εισοδήματος και Συνθηκών Διαβίωσης των Νοικοκυριών, της Ελληνικής Στατιστικής Αρχής (ΕΛΣΤΑΤ), η συγκεκριμένη διπλωματική εργασία επιχειρεί να εντοπίσει τα νοικοκυριά που δεσμεύονται από κάποιου είδους κόστους στέγασης κατά τη χρονική περίοδο 2007-2022, δίνοντας έμφαση στις επιδράσεις πολλών παραγόντων στη πιθανότητα αθέτησης αποπληρωμής στεγαστικού δανείου ή ενοικίου. Η μελέτη αυτή αναγνωρίζει τη σημασία του κινδύνου αθέτησης αλλά και το ευρύτερο φάσμα των οικονομικών προκλήσεων που ενδέχεται να αντιμετωπίσουν τα νοικοκυριά, ιδιαίτερα στον τομέα των δαπανών που σχετίζονται με τη στέγαση. Δεδομένου των αποτελεσμάτων και έχοντας λάβει υπόψιν δημογραφικά χαρακτηριστικά, χαρακτηριστικά σε επίπεδο νοικοκυριού καθώς και περιφερειακά στοιχεία, διαπιστώνουμε ότι παρατηρούνται χαμηλότερα επίπεδα αθέτησης υποχρεώσεων για νοικοκυριά με υψηλό εισόδημα. Κατά συνέπεια, αυτή η μελέτη έχει ιδιαίτερη σημασία για την πρόληψη και τον έλεγχο των νοικοκυριών με κόστος στέγασης.

Λέξεις κλειδιά: Ελλάδα, Νοικοκυριά, Στεγαστικό δάνειο, Αποπληρωμή ενοικίου, Πιθανότητα αθέτησης

Contents

1 Introduction **1**

2 Literature Review **5**

2.1 Literature Review 5

3 Data **7**

3.1 Data 7

4 Econometric Modelling **26**

4.1 Econometric Modelling 26

5 Empirical Results **30**

5.1 Empirical Results 30

5.1.1 Determinants of constraint households 30

5.1.2 Determinants of housing cost and hardship repayments . . . 36

5.1.3 Determinants of default in housing 44

6 Conclusions **49**

References **51**

List of Tables

1 Summary statistics: Mean values of key variables 18

2 Summary statistics (households over a 4-year period) 23

3 Relative risk ratios on each type of constraint household 35

4 Average marginal effects on the likelihood of having housing cost . . 39

5 Average marginal effects on the likelihood of experiencing hardship with loan or rent repayments 43

6 Probabilities of default estimates 48

List of Figures

1 Distribution of Constraints households 12
2 Housing cost 13
3 Difficulties in paying loan or rent 14
4 Default rates (households over a 4-year period) 25

Chapter 1

Introduction

Initially, a conceptual framework is outlined where household finance studies the ways in which households use financial instruments to attain their objectives. The field has grown rapidly in recent years, with considerable emphasis on nonfinancial assets that are often financed with debt. Housing is the most important asset of the household sector and for many people rents or loan repayments make up a relatively large proportion of their , while mortgages represent both the potential and complexity inherent in the international comparative study of household finance. As a variety of factors influence housing markets, our study mostly focuses on those that might have a bearing on the repayment behaviour of home loan borrowers or tenants at the level of individuals and households. So, taking all the above into consideration and motivated by the above statements, I examine the probability of default, involving both tenants and borrowers, in Greece with the use of annual household and individual-level data, drawn from the Greek Survey (EU-SILC) for a period of 16 years (2007-2022), published by the Hellenic Statistical Authority (ELSTAT). However, while the study was concerned with the features that result in default in payment of housing loan instalments or rents, in light of our study, it was also investigated and then reported the findings in relation to the probability of having housing cost and being in difficulty with their mortgage or rent payments, respectively. This broader perspective could provide

insights into the factors that contribute to financial stress within the context of housing, distinguishes itself from the previous studies. In continuing analyzed all of them, elaborately.

As the rising incidence of housing debt amongst both owner-occupiers and tenants and the main causes of rent and mortgages arrears were the research incentives, the primary step, to initiate this investigation, includes the analysis of the collected data in order to identify borrowers who are encountering challenges with debt repayment and tenants who are struggling with rent payments. Proceeding with deeper analyses, we examine the underlying factors contributing to households' financial struggles and estimate the probability of default on repaying loan or rent. The micro-level analysis is going to be based on several multiple regression models that are going to be estimated with the help of the statistical analysis software, Stata. Firstly, we estimate the probability of a household belonging separately in each category using a multinomial model and secondly conduct an analysis of the relationship between the likelihood of having housing cost and the likelihood of struggling to keep up their mortgage or rent repayments, using two logit models, respectively. Finally, we estimate a linear regression model in order to examine the changes on the probability of debts that are left outstanding.

To shed light on the factors that have a bearing on the repayment behaviour of home loan borrowers or tenants, empirical research is carried out encompassing a sample of households in the period from 2007 to 2022. Firstly, the empirical evidence obtained lets us conclude that older age is associated with a lower probability of incurring housing costs. But we also report results indicate that gender and marital status have a negligible impact on default probability, in terms of the households with housing costs. My investigation also provides quantitative evidence for one of the core determinants of the level of default probability, the level of education. More particularly, I find that the outcomes of delinquent loans

are significantly and negatively related to the university attainment. Evidence show that default probability differs depending on the economic activity of the individual. Housing cost increasing more at the lower ends of income, with these groups might be more susceptible to economic fluctuations, impacting their repayment capacity. Also, household size could eventually fortify the probability of struggling with mortgage or rent repayments. The relationship between areas and the probability of mortgage default is multifaceted and varies across regions. Nevertheless, lack of understanding about mortgage terms and interest rates, due to the limited financial literacy can potentially elevate the risk of default.

The current study findings have important implications for those programs aimed at narrowing the gap between struggling households and the probability of default, which is the framework for coordinating economic policies. Generally, striking a careful balance between financial stability and productive risk-taking by the commercial banks in Greece remains a continuing challenge. Based on case studies, government policies that target the issue of excessive household debt facilitate and they can be instrumental in preventing prolonged economic contractions. Armed with accurate risk assessments lenders can make informed decisions regarding loan origination, underwriting, and pricing. Moreover, regulators can better shape guidelines that promote responsible lending practices. So, by fostering a financially secure and confident population, governments can significantly reduce the number of households defaults and as a result, reduce the negative consequences of default events improving housing quality and contributing to the overall stability of the housing market and financial system. As such, the residential mortgage default literature stands as a crucial pillar in the broader landscape of financial research.

To get an idea about, the remainder of the present dissertation's structure is organized as follows. Chapter 2 provides a background from the existing lit-

erature on the concept of household finance, covering prevailing frameworks for analysis and main findings related to our investigation. Chapter 3 refers to the data sources, considers descriptive statistics and graphics are used to display data. Methodological issues and empirical models are presented in Chapter 4. In addition, Chapter 5 shows and discusses the main estimation results of the analysis to shed light on the factors that impact their ability to make regular mortgage payments. We conclude with a brief summary of key points in the Chapter 6, which is followed by the references.

Chapter 2

Literature Review

2.1 Literature Review

The residential mortgage default literature is fundamentally driven by a quest for a precise comprehension of the risk of default as conveyed by loan and borrower attributes. This desire for accuracy stems from the critical need to effectively assess and predict the likelihood of borrowers defaulting on their residential mortgage loans. So, this area of research is dedicated to unraveling the intricate relationship between various characteristics associated with both loans and borrowers.

In order to provide possible explanations for our findings we review the relevant literature. The empirical work of Agarwal et al. (2011) explores the risks of mortgage default and prepayment specifically within the context of moderate- and low-income households. The authors likely investigate the factors that contribute to the likelihood of mortgage default and prepayment among individuals with relatively lower incomes. The study funded by Bucks and Pence (2008), looking at housing debt, examined the relationship between mortgage literacy and default. They found that households who were less informed about their mortgage terms and conditions were more likely to default. This points to the importance of understanding individual financial literacy as a determinant of default. One of the

most influential studies in this topic is the work of Mian and Sufi (2009) who demonstrated that an increase in the supply of mortgage credit, driven by securitization, led to a boom in housing demand and subsequently a spike in defaults. Personal characteristics, including income and educational attainment, were found to influence the likelihood of default during the crisis.

Campbell et al. (2011) attempts to document the impact of forced home sales (due to default) on house prices. The authors found that while individual financial characteristics like income and employment status play a role in defaults, neighborhood effects also matter. A higher number of foreclosures in a particular neighborhood can depress home prices in the area, leading to further financial distress for homeowners. Also, the study of Gerardi et al. (2013) demonstrated that individual numerical ability is a significant predictor of mortgage default, suggesting that cognitive abilities can play a significant role in household financial decision-making. Another paper focused on the UK and found that individuals with low levels of financial literacy are significantly more likely to default on their debts (Disney et. all.,2015). However, they also found that low financial literacy households who seek credit counseling are less likely to default than similar households who do not seek counseling.

So, the phenomenon of housing loan defaults and the possible drivers of the defaults have been extensively researched and reported in the literature, and it continues to attract the attention of researchers. There is an extensive and growing body of literature in the arena of mortgage default, some of which form the basis of the present paper. It is worth noting, our empirical results are broadly consistent with the more general findings in the literature, as appears in the section of the estimation results.

Chapter 3

Data

3.1 Data

In this part of my analysis, I present the description of the data used in the specific research. Particularly dataset, documenting household financial behavior, was obtained from the European Union Statistics on Income and Living Conditions (EU-SILC) survey. So, our dataset concern households in Greece and the time period that was covered using these data was the years 2007 to 2022. This survey is the only source of data on income, living conditions and some aspects of health and health seeking behaviors at household-level and individual-level (for persons aged 16 and more). So, we provided imports data from an Excel file for each year that contains multiple sheets (H-FILE, D0FILE, R-FILE, P-FILE). Firstly, H-FILE data contains extensive information on household characteristics such as the unique household identifier, the disposable income, the type of housing tenure etc. Turning to D-FILE, demographic or area-level information are contained. R-FILE data includes individual records and separate characteristics like the unique personal identifier, gender, and age-related information. Last but not least, P-FILE pertain to personal characteristics or statuses of individuals. In particular, the selected data of this file included In particular, the selected data of this file included the working-age individuals or those aspiring to be part of the

economically active population, aged 16+, that were asked to display information about their marital status, educational level, with the inclusion of data to their employment status.

Let's give a brief description of the definitions of variables. Specifically, several variables will be used including the year of the survey, household identification number (HID), total disposable household income (without inflation factor), dwelling type, tenure status, number of rooms available to household, current rent related to occupied dwelling (if any), mortgage principal repayment (monthly principal repayment mortgage). The study also considers the variable that indicates whether the household has been in arrears on mortgage or rent payments in last 12 months, net interest repayments on mortgage and gross interest paid on mortgage. It also includes a hierarchical classification of geographical regions, the degree of urbanizations specifying whether a location is urban, semi-urban and rural. Further, variables denoted the year of birth and the gender of the individual, the marital status, the year the individual graduated, the highest level of education attained (i.e. primary education, secondary education, post-secondary education, tertiary education, and Ph.D.) and three employment status groups (i.e employed, unemployed, not economically active).

In terms of data pre-processing, as appending H-FILES, D-FILES, R-FILES and D-FILES in four Stata Format (.dta) Data Files, respectively, we calculate the average age, which is measured in years from the birth of the person to the time of the survey, the average share of females, and the average value of the household size, using the R-FILES. We, also, redefined several variables like "married" based on the values of the variable that represent the marital status (PB190), "university" based on the variable of the highest educational attainment (PE040 for the years of 2007-2020 and PE041 for the years of 2021 and 2022) and display whether an individual graduated from university, "employed", "unemployed"

and "inactive" variables are also contained, based on the variable of activity status (PL030 for the years of 2007 and 2008, PL031 for the years of 2009-2020, PL032 for the years of 2021 and 2022), in P-FILES. Similarly in D-FILES, we created a series of dummy variables like "urban", "semiurban" and "rural" based on the values of the variable that represent the degree of urbanization (DB100), "nuts1_D" based on the variable of regional categorization: "nuts2" (for the years of 2018-2022) and "nuts1" (for the years of 2007-1017) combined with the year. So, four regional dummy variables are generated, refers to Attica, Crete and Islands, Northern Greece, and Central Greece. Thence, as calculating the average values of the new variables in both files and saving the new files, we merge the above datasets using the personal ID (PB030) and the year of the survey (DB010) as keys. Only observations that matched in all files are kept.

For our analysis, we create a dummy variable based on tenure choices of the households. More specifically, the tenure of a household indicates whether household in a private dwelling or in renting housing (HH020 for the years of 2007-2009 and HH021 for the years of 2010-2022). Those cases in which accommodation is provided rent-free are omitted from the analysis. Initially, we need to calculate the total housing cost per room. So, to do this, we take the monthly cost of renting their living space (HH060) or the net interest paid on mortgage (HY100G) and divide it by the number of rooms (HH030), for renter-occupied households and for households with mortgage, respectively. Now, as calculating the total expenditure for both renters and owners with a loan, we are able to define whether a household has expenditure occupied the dwelling or not. So, at this point we create two categorical variables, "rent_roomD" and "loanmort_roomD" which take the value 0 if the rent to room ratio and the mortgage payment to room ratio, respectively, is missing or equals to zero and the value 1 if the value of index is positive.

While the first objective of this study was to identify households struggling with

mortgage or rent repayments, we create another dummy variable which indicates if a household has been in arrears in the past two 12 months or has not. So, we are in a position to construct indicators at household level on whether the financial obligation related to housing debt or rent is satisfied or not. So, the usefulness of the above variables is exploited in the construction of a new variable that grouped into domains:

1. *Households without arrears on mortgage or rent payments.*
   a. Households that living in owner-occupied housing and do not deal with arrears of payment of mortgages.
   b. Tenants were not experiencing a problem with payment of rent.
2. *Households with arrears on mortgage or rent payments.*
   c. Households that living in owner-occupied housing in arrears of payment of mortgages.
   d. Tenants were experiencing a problem with payment of rent.

So, there is an interesting distinction in the owner-occupied and rent-occupied households with or without arrears on the different types of their housing cost. An illustration of these outcomes is also depicted in Figure 1. Concentrating on the variable "mycategory", we draw our attention to Figure 1 to get a first flavor of constraints households. The time frame is covered by the X-axis. More specifically, the X-axis is labeled with consecutive years from 2007 to 2022. The Y-axis represents the "Share of each Type of households Facing Constraints". The legend of the bottom of the graph explains the marker corresponding to each category. The first group includes owners without arrears, the second one those who were tenants without arrears, the third one is about owners with arrears, the fourth one refers to tenants with arrears, while the fifth one contains those who were owners without housing cost. As we observe how they change over time, lots of fluctuations occurred with the most notable changes being during the crisis period. In particular, the range of the rate of households (owners or tenants)

without payment-related problems lays between 0.05 and 0.15, while the rates of households with payment-related problems lays between 0.01 and 0.08, across the observation period.

In terms of constraint households, a significant drop is displayed for those who are tenants without arrears and those who are owners without arrears for the period 2009-2013 and the period 2008-2015 which encapsulate the global financial crisis, respectively. This longer timeframe for owners might reflect the more extended financial commitments tied to mortgages and property ownership. In contrast, we clearly notice a rising trend concerning the proportion of both tenants and owners with arrears about the same period. So, this complements the earlier observation, as fewer households without arrears would mean more with arrears. Yet not only the gap between owners with and without arrears and similarly between tenants with and without arrears did it diminish in these years, but also the divergence was counterbalanced since the increase in the percentage of those who are owners or tenants with arrears had gradually converged to the downward trends of the other categories (Owners and Tenants without arrears). To rephrase it, the simultaneous decline of households without arrears and the rise of those with arrears highlight the broad economic strain felt by the population during this period. So, this data underscores the vulnerability of both renting and owning households to economic downturns, which lead to job losses, reduced income, and other financial challenges, making it more difficult for households to meet their housing cost obligations. As the above are graphed in Figure 1, we can also see that tenants without arrears are the ones with the higher rate constantly, as well as the one was affected the most by the economic crisis. Nonetheless, the particular rate was steadily higher than the others.

As the Greek economy overcame the recession in 2015, a noteworthy increase is presented for those without arrears and a remarkable percentage fall for those with arrears up to 2019. To be more accurate, as the economy improved, more

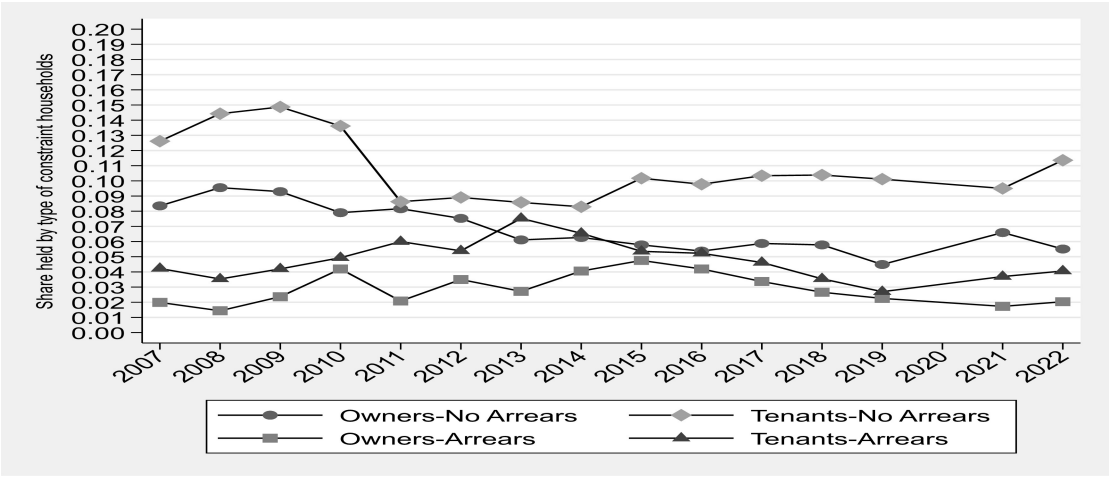


Figure 1: Distribution of Constraints households

households were able to meet their housing cost obligations, reducing the number of households in arrears. However, from 2019 onwards, there was a spike of all the categories, except for Owners with arrears, that was the result of the pandemic of coronavirus (Covid-19) and the consequent lockdowns, which strained many households. The fact that in the Category of owners with arrears is not reflected the unprecedented global economic challenges that brought the Covid-19 pandemic might be explained by the homeowners which have a more extended financial commitment and better financial buffers or were more conservative in their financial decisions, making them less immediately vulnerable to short-term economic shocks. In addition to this, with global interest rates falling during this period, some homeowners might have taken advantage of refinancing opportunities, reducing their monthly financial burdens. Undoubtedly, the distinctions between owners and tenants in response to economic shifts reflect the different dynamics and challenges of the ownership and rental markets.

Two indices (index0 and index1), also, are created. The first one differentiates households with and without housing costs, while the second one differentiates households with and without arrears. The following graph (Figure 2) depicts the distribution of households with housing cost. It shows the fraction of housing units with housing cost that are owner occupied on the left axis and a 15-year period on

the right axis. The line of the slight and stable downward trend during the period of interest, implies that the decline in the proportion of overall level of mortgage holders and tenants was uninterrupted across time. To be more accurate, it formed more evenly and remains in the same level over the sub-period 2010 to 2014 and then declined steadily between the years of 2015 and 2019, as seems to be unable to meet obligations and stood at 21% in 2022. More analytically, the aftermath of the economic crisis leading to a slowdown in real estate transactions. Fewer people taking up new mortgages or renting homes due to economic uncertainty.

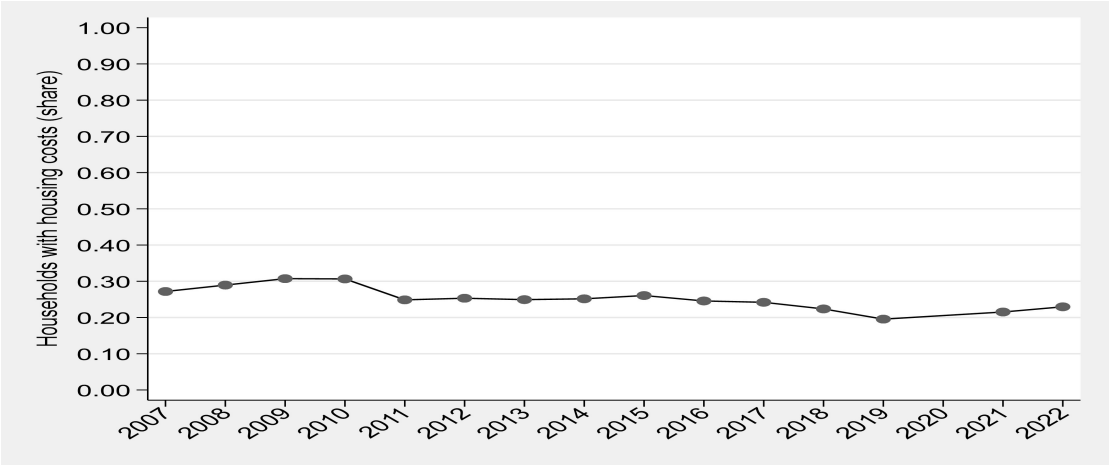


Figure 2: Housing cost

So, figure 3 plots the average share of households with payment-related problems. In the year 2008, we observe that difficulties with homeowner loans have an average increasing rate which had reached unprecedented levels, in early 2014. Specifically, over two-fifths (42%) of all households with housing cost were in arrears of payments. Most likely, the main reason for this inability to comply with housing-related obligations can be found in the worsening of the Greek economy due to financial crisis which started with a subprime mortgage lending crisis in 2007 and expanded into a global banking crisis in mid-2008. During this period, many households took on high levels of debt. In particular, this rapid accumulation of household debt made households more vulnerable to financial shocks and economic downturns. Moreover, as looking at Figure 3, the share of households

facing payment-related issues has been on a declining trajectory toward the end of 2013 and into 2014 for an extended period of time (post-recession period), where global economic recovery, with markets stabilizing is happened, while it has yet to reach its pre-crisis levels, at started period, indicating the long-lasting impacts of the recession. Apparently, there is a noteworthy percentage reduction until the first effects of the emergence of pandemic became visible. So, with respect to the period of 2019-2021, Covid-19, in turn, contributed to the fact that households with arrears remain quite stable.

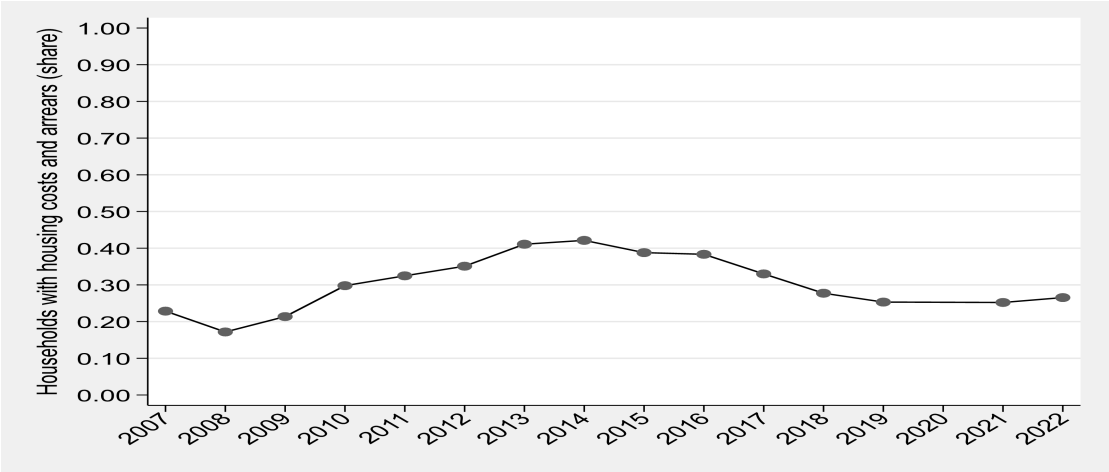


Figure 3: Difficulties in paying loan or rent

In this study, we will examine four models which are explained in the next section. The dependent variable we use in model (1) is "mycategory" consisting of 160,348 observations separated into 5 subcategories. Descriptive analysis of the sample's data constitutes a basic tool for the correct and impartial interpretation of the results. So, Table 1 provides a visual map for the data collection. The number of observations and the average share of the variables to each category are presented. So, comparing the average allocations of household among five categories we mentioned before, we can see that the vast majority of population is living in a household owning their home, with an ownership rate of about 85 percent (Category [1], Category [3] and Category [5]) , while the remaining (One in ten) is living in rented housing (Category [2] and Category [4]). According to

the second row of Table 1 which shows the percentage shares of population configuration, those that live in a housing unit and do not have a loan headed 75% of Greece households (Category [5]). Now, the Category [1] and the Category [2] regards the households with a housing cost (mortgage or rent) which are able to pay for. These groups form 6% and 10% of the questioned population of the survey, respectively. While it is clear that cases referred to Greece households (landlords and tenants) in arrears of payment of mortgage or rent has been reaching 3% and 4,5% of our sample, respectively (Category [3] and Category [4]).

Key takeaways from the household characteristics (Panel B) indicate that on average, the age of owners with and without arrears is around 40 years and 42 years, respectively, while the age of owners who do not have to pay mortgage is around 58 years. This implies that younger owners may face financial challenges or uncertainties that could lead to payment delays or difficulties, while the older owners, potentially with more life experience and financial stability, might be more diligent in ensuring their debt, if any, have been paid on time. Turning to the borrowers with and without arrears, we observe that their average age fluctuates at the same levels with owners, respectively. Especially, borrowers in Category [2] have an average age of 39 years and borrowers in Category [4] have an average age of 39 years. So, age seems to play a significant role in the likelihood of having cost but might not be as significant factor for their responsibility in repaying borrowed money or rent. Gender distribution remains reasonably balanced, with female representation ranging from 50.2% to 55.0% across the categories. Notably, households in category [1] exhibit the highest proportion of married individuals at 68.8%, whereas category [2] showcases a noticeable lower percentage at 44.2%. Case in point, something worth pointing out, is that the owning property or assets may be more feasible for married couples due to combined incomes, shared financial responsibilities, or other economic benefits of marriage, while borrowing might be more common among younger individuals or those at earlier life stages

who may not yet be married.

Furthermore, we studied how people in Greece were distributed among the examined categories based on the educational institution of the highest education they have completed, when it refers to the university level. Specifically, the highest percentage of the population has a master's or doctorate degree refers to the first and second category. To put it another way, people who have high levels of education belongs to this percentage of the total population without arrears. Generally, advanced degrees correlate with higher earning potential and more stable employment professions with tenure, reducing financial unpredictability. So, owners with high education levels suggests that may have leveraged their education to secure better-paying jobs, enabling them to purchase and maintain ownership of assets without falling into arrears, while for borrowers, the absence of arrears indicates good financial management, as a result of a higher education that leads to these individuals borrowed for specific purposes (e.g., further education, business investments) and had a clear plan or means to repay without falling behind. Moreover, we are going to study the distribution of the employed, unemployed, and inactive citizens in Greece across the five different levels of their living conditions. So, the employment data reveal significant employment in categories [1] to [4], with the most substantial employment in category [1] at 56.6%. In contrast, unemployment appears highest in category [4] at 20.2%. The average household size tends to be larger for categories [1] and [3]. This fact can be explained by owners either those that are economically well-off enabling them to afford larger homes and maintain larger families (Category [1]) or those with multi-generational households or extended families which live together (Category [3]).

Other interesting facts emerge from this table, income data highlights categories [1] and [2] with higher average annual income, implying that they enjoy better financial stability and possibly access to more resources and opportuni-

ties. In addition to this, it is important to see the percentage of individuals that reside in each region of Greece. As we can see, the regional breakdown showcases varied household distributions across regions like Attica, Crete and Islands, Northern Greece, Central Greece, with distinct urban, semi-urban, and rural categorization. As far as islands, are concerned the smallest percentage across all the categories may be due to lower population densities. To end up, the share of the total population living in urban areas ranged from 32.6% to almost 52.8%, with the highest percentages are observed in the categories related to homeowners.

Table 1. *Summary statistics: Mean values of key variables*

Panel A	Total	[1]	[2]	[3]	[4]	[5]
Observations	160,348	10,104	16,868	4,809	7,268	121,299
(% of households)	1.000	.063	.105	0.030	.045	.756
Panel B						
Household Characteristics:						
Age	53.939	41.933	38.785	40.525	37.664	58.553
Females (%)	.539	.506	.502	.512	.503	.550
Married (%)	.549	.688	.442	.658	.463	.552
High education (%)	.193	.347	.298	.206	.171	.167
Employed (%)	.346	.566	.543	.495	.461	.288
Unemployed (%)	.083	.081	.093	.177	.202	.071
Inactive (%)	.569	.352	.363	.327	.336	.640
Household size	2.393	3.094	2.241	3.182	2.471	2.320
Net annual income	16,075	25,196	16,174	15,983	11,013	15,608
Number of rooms	3.173	3.575	2.675	3.466	2.595	3.232
Geographical Groups:						
Attica	.259	.337	.366	.251	.377	.231
Crete and Islands	.154	.127	.147	.123	.164	.158
Northern Greece	.268	.348	.28	.366	.253	.321
Central Greece	.268	.187	.197	.258	.204	.288
Population Density Domain:						
Urban	.326	.389	.528	.329	.523	.281
Semi-urban	.230	.281	.279	.270	.252	.216
Rural	.443	.329	.191	.400	.224	.502

Source: Income and Living Conditions (EU-SILC, 2007-2022), Hellenic Statistical Authority (EL.STAT.), Public Use Files (Microdata)

Notes: [1] Owner with housing cost without arrears, [2] Tenant with housing cost without arrears, [3] Owner with housing cost with arrears,[4] Tenant with housing cost with arrears, [5] Owner without housing cost.

In this stage of our analysis, special attention was given in processing the data as well there are differences in the number of observations regarding each household, and these observations correspond to different points in historical time. So, we create a balanced panel data set of longitudinal households observed periodically over a 4-year period, to be analyzed. The use of panel data allows us to consider changes over time within the same households, which can provide a more nuanced understanding of the relationships that indicated by the analysis of the variables.

There are two categories of loan status: default or not default. The first one pertains to households which found it harder to meet their mortgage or rent repayments. Most commonly, mortgage and rent arrears were a result of a multiple debt problem. In our investigation, depending on households' payment history and whether their difficulties are long or short term, we decide who is more likely to be behind on their mortgage or rent repayments. So, in this study, the levels of debt were measured by reference to the time scale over which the debt has accrued, namely the number of missed yearly mortgage or rent repayments that were equivalent to the arrears. To be more accurate, firstly, we define the housing status of each household for the 4-year period, getting lags for previous years. So, since the time sequence of the housing status is derived, we calculate the probability of default, dividing the years of outstanding payments by the period with housing cost.

Table 2 provides a breakdown of both full and selected, with those with housing costs, sample, respectively. We compute averages across households using the variables are included in Table 1. Firstly, as it can be seen in Table below, we can extract that the average level of tenants' or borrowers' probability default varied according to the sample. More specifically, all Greek households has a probability of 8.5% being indebted, whereas for the selected sample, the mean value of the probability of default is equal to 0.335, which implies that if the homeowner is

borrowing constrained or being a tenant has a considerably higher probability at 33.5%. This suggests that having housing costs is associated with a significantly higher likelihood of default. Looking at the second row of Table 2, the average age concerning all households is 55.5 years old, while people on households with housing costs aged, on average, roughly 42 years old. So, as younger ages are associated with family formation, the phase of acquiring their first home often involves taking on a mortgage or paying rent, hence they are more likely to incur housing costs. The role of gender doesn't seem to drastically differ between the two groups. Across all households, 54.5% are female-led, whereas for those with housing costs, the percentage drops slightly to 50.6%. Moreover, about one half of the whole sample is mentioned to married people, compared to a slightly higher proportion of 56.9% among those with housing costs, suggesting that marital status might not be a major determinant for incurring housing costs. Nevertheless, it is well accepted that married couples are more likely to have dual incomes, which might make it easier to afford housing costs such as mortgages or rent.

In addition, the collected data as illustrated in Table 2, provides information on the education level. This can detect possible differences for those households that having a higher education degree between full and selected sample. Observed that, only 19.2% of all households have a higher education degree, while this increases to 28.5% for households with housing costs. So, higher level of education is associated with better job opportunities and higher salaries, which can make housing more affordable for these individuals. As we can see, studying the data, the mean of the employed variable for all households and households with housing costs equals to 0.326 and 0.513, meaning that the 32.6% and 51.3% of the individuals included in this variable are employed, respectively. Except for the category of those who participate in labour force, 8.5% of all households are unemployed, rising to 12.6% among those with housing costs. So, having a job appears to correlate with having housing costs, but unemployment is also slightly more prevalent

among those with housing costs. This complex relationship between employment status and housing costs provides valuable insights into the financial vulnerabilities and complexities faced by different groups and as a result, it underscores the importance of social safety nets and financial planning in maintaining housing stability.

The average household size is about 2 individuals for all households, but the corresponding number for household members among those with housing cost is about 3 individuals. Larger households often consist of families with children and the presence of them usually necessitates a larger living space, and hence, potentially, more housing costs. In the instance of income, for all households, the annual disposable income is €15,695.71, whereas the income of the selected sample corresponds to €17,734.05, on average. Higher income seems to be associated with the likelihood of having housing costs. The average number of rooms is almost similar in both categories but slightly lesser for households with housing costs. Numerically speaking, 3 rooms is the mean count of rooms in both set of dwellings. With regard to regional dummies, looking at the rows of Table 2 which allocate households who resided on each of the four regions separately, we clearly ascertain that, on average, 33.2% of the total households being recorded in the households based in the north of the country followed by the central while the lowest rates belong to Crete and Islands. Furthermore, as far as the location of the people that having a housing cost, we can see in the below table, that the majority of households lives in the region of Attica, something that is to be expected given that it is the most populated region that also includes the capital of Greece, namely Athens. The second highest percentage refers to the people who live in Northern Greece, with those people accounting for 32.1% of households with housing cost. As far as the islands are concerned and the average number of those who lived in, our results state that they have the lowest participation, as also demonstrated in the case of all households.

The share of the total population living in urban areas in Greece was on average 31.6%, compared to those who lived in the same areas, in terms of the selected sample, which consists of the substantial fraction of 45%. The corresponding share of population in semi-urban areas is on average only about one-third for both samples. Lastly, the average number of all households and households with cost who lived in rural areas are 48% and 30.3%, respectively. Instead, urban areas tend to have a greater proportion of households that incur housing costs, likely due to factors such as higher property values, increased demand for housing, and a greater availability of housing options like apartments and condominiums that often come with associated costs like rent or mortgage payments. On the other hand, rural areas generally have a lower percentage of households with housing costs. This could be attributed to various factors such as lower property values, a higher prevalence of home ownership without mortgages, or the existence of generational homes that have been passed down and thus do not have associated mortgage costs. Rural households may also benefit from owning larger parcels of land, potentially with multiple dwellings, that can be used to offset living expenses. The analysis of all households is based on 19,925 observations, while the subset with housing costs has 5,051 observations. Thus, in Table 2 below, the results of said percentages are presented.

Table 2. *Summary statistics (households over a 4-year period)*

	All households	Households with housing cost
Probability of default	.085	.335
Average age	55.531	41.692
Share females	.545	.506
Share married	.552	.569
Share higher education degree	.192	.285
Share employed	.326	.513
Share unemployed	.085	.126
Share inactive	.588	.360
Average household size	2.413	2.744
Average net annual income	15695.71	17734.05
Average number of rooms	3.191	3.103
Attica	.243	.326
Crete and Islands	.144	.135
Northern Greece	.332	.321
Central Greece	.278	.217
Urban	.316	.450
Semi-urban	.202	.246
Rural	.480	.303
Observations	19,925	5,051

Source: Income and Living Conditions (EU-SILC, 2007-2022), Hellenic Statistical Authority (EL.STAT.), Public Use Files (Microdata)

Figure 4 illustrates how the likelihood of failing to meet housing-related financial obligations has evolved over time for two separate categories: the entire pool of households and the subgroup that has housing expenses. We clearly notice a significantly percentage rise on the default probability concerning the households with housing cost, for an extended period of 2008-2014, that is originating from the financial crisis. The spike in default risk among these households suggests that the financial crisis had a disproportionate effect on those who were already

financially burdened by housing costs. These problems appear to be concentrated among those least able to bear the consequences of crisis. In other words, the crisis intensified the struggles of those least equipped to cope with economic shocks, pushing the default rate to a concerning 45% by 2014. Even though, probability had soared to the highest peak by 2014, highlighting the immense financial pressure these households faced, it tapered off, reaching roughly 30% (although it's still a substantial proportion) of the selected sample in 2019, possibly indicating some level of recovery or adaptation.

Contrastingly, for the overall population, as depicted in Figure 4, it is obvious that defines a stable trend over the same time period which implies that the proportion of default rates was approximately uninterrupted across time. The contrasting trends between these two groups imply that households with housing costs were much more affected by the financial crisis compared to the general population. While the broader population maintained a relatively steady default rate, households with housing costs experienced substantial fluctuations, largely due to the economic hardships amplified by the crisis. Therefore, the following Figure serves as a visual representation of how different segments of the population are differently affected by economic shocks and can be a valuable tool for policymakers to understand where to focus their interventions.

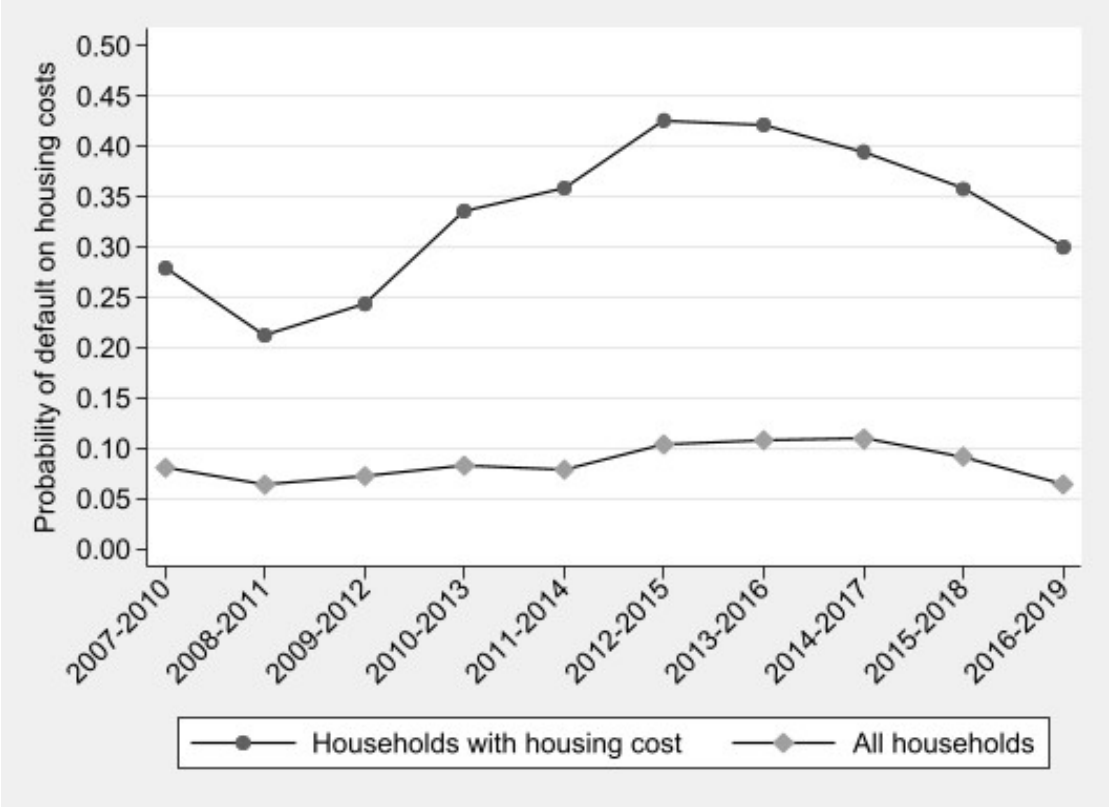


Figure 4: Default rates (households over a 4-year period)

Chapter 4

Econometric Modelling

4.1 Econometric Modelling

Carrying on with the empirical analysis, we first present the methodology framework. So, before presenting our empirical results, it is necessary to delineate the econometric models that form the baseline for estimating the effects of the predictor variables on the dependent variable that we examined. More specifically, in this section, methodological guidelines describe the 4 models that we would estimate. These models have been chosen based on their suitability to the nature of the data and the research questions we aim to address.

In our thesis, the analysis that follows is divided into two parts. In the first part, we use cross-sectional data over certain period (2007-2022) and apply a multinomial logit model to demonstrate the effect of individual and household characteristics on the probability of the outcomes in Y, as they mentioned in the description of the variable "mycategory" and thence, we apply two logit models to examine the effects of the same determinants on the probabilities of housing cost and hardship repayments. And in the second part of our analysis, special attention was given in processing the data as well there are differences in the number of observations regarding each household, and these observations correspond to

different points in historical time. So, we create a balanced panel data set of longitudinal households observed periodically over a 4-year period, to be analyzed. In this instance, of the analysis using data over time, we implement a regression using the OLS method in order to uncover the key factors that significantly influence the default probabilities.

An empirical model (multinomial logit) that investigates the effects of the explanatory variables across different household type categories is specified by:

$$P(Y_i^k) = \alpha + \beta^k * X_i^k + \epsilon_i \tag{4.1}$$

where $P(Y_i^k)$ denotes the probability that the household i belongs to the group k where k takes five values that correspond to the cases "Owners-No Arrears" ($k = 1$), "Tenants-No Arrears" ($k = 2$), "Owners-Arrears" ($k = 3$), "Tenants-Arrears" ($k = 4$) and "Owners-No Housing Cost" ($k = 5$), X_i is a vector of explanatory variables that are expected to influence the outcome of interest, and ϵ_i is the error term capturing unobserved factors that may influence the dependent variable. The vector X_i consists of the variable "female" which denotes the value of a categorical outcome variable and especially, getting the value one when the gender is female, and zero when a person is male, the variable "Married", a variable showing if a person is married, the variable "university", a dummy getting the value one when a person has been to university, the variables "Employed", "Unemployed", depending on whether the individual is employed or unemployed, the variable "Household Size" and "Number of rooms", variables that measuring the number of individuals and the number of rooms in each household, respectively, the variable "Income Quantiles", a categorical variable that takes values from 1 to 10 depending on household income level, the variables "Crete and Islands", "Northern Greece", "Central Greece" indicating the region of residence of the respondents. Also, we included among the regressors, three dummy variables indicating if a person is living in urban, semi-urban or rural areas.

Afterwards, applying a logit model, we examine which sectors at individual and household-level are more probable to affect the probability of a household having housing cost. This same reasoning, a clear understanding of the relation that subsist between the probability of a household having struggled to afford housing cost and the same explanatory variables as before, is achieved through a second logit model. The empirical function generated is the following:

$$P(Y_i = 1) = \alpha + \beta * X_i + \epsilon_i \tag{4.2}$$

In the first logistic regression, the dependent variable is a dummy indicator that denotes whether a household deals with housing cost, while in the second one, the dependent variable is a dummy indicator that denotes whether a household falls into arrears. The set of independent variables included in this model are those included in the previous model. ϵ_i is the error term capturing unobserved factors that may influence the dependent variable.

Finally, we estimate two separate regression models with the same form, one concerning all households and another one concerning only households with housing cost. In order to investigate the impact of some of the basic households' and individual's characteristics on the dependent variable which demonstrates the likelihood of default in Greece during 2007-2022, an econometric model has been developed and a multiple regression model is used. The linear regression model that will be used to estimate the determinants of housing default is the following:

$$Y_i = \alpha + \beta * X_i + \epsilon_i \tag{4.3}$$

where Y_i denotes the likelihood of default for the household i, X_i is a vector of explanatory variables that are expected to influence the outcome of interest, and ϵ_i is the error term capturing unobserved factors that may influence the depen-

dent variable. The set of independent variables included in this model are those included in the previous model. This model is estimated for the entire sample of household and separately for households with housing cost.

Chapter 5

Empirical Results

5.1 Empirical Results

5.1.1 Determinants of constraint households

Table 3 that follows, summarizes the changes in the probability of the household belonging to category [1], category [2], category [3] and category [4] in contrast with the category [5]. So, the reference category defined in this case is the category [5]. In other words, we examine the probability of a household having a housing cost, depending on a variety of socio-economic variables, including age, gender, marital status, educational attainment, employment status, household size, income quantiles, number of rooms in the dwelling, and geographical location. Initially, it is significant to mention that the estimated coefficients transformed to relative-risk ratio and computed at the mean of each independent variable.

From a preliminary glance at the top half of the table, older age is consistently associated with a slightly lower probability across all household types to belong to each of them, separately, compared to those that have not housing cost (Category [5]). The coefficients ranging from .926 to .962, all statistically significant at the 1% level. This outcome is consistent with our expectations and the immense literature surrounding the household finance and economics. According to the

well-accepted economic theory, individuals plan their consumption and savings behavior over their lifetime, as they save and accumulate wealth during working years and dissave in retirement (Modigliani and Brumberg, 1954). So, by the time they are older, they are more likely to have completed substantial financial commitments like housing loans. Female heads of households tend to have a higher risk of having cost, especially among homeowners (both with and without arrears), while the impact is negligible among tenants. Women's lower earnings, on average, can result in housing costs as a larger portion of their income, affecting their financial stability and make housing costs more burdensome. This result coincides with previous studies (Blau and Kahn, 2017).

Looking at the estimates of the third row, we detect that marriage appears to increase probability of housing cost for owners, but its effect on tenants is minimal. The explanation of these findings is fully compatible with the findings of Browning and Lusardi (1996), as well as married couples can share costs, they may view homeownership as a long-term investment and may be willing to take on higher costs as a result. On contrast, housing policy literature often discusses rent controls and housing benefits that can shield renters from excessive costs. Tenants often have more variability in their incomes and may opt for housing solutions that are more adaptable to their financial situation the present, mitigating the impact of marriage on housing costs (Desmond and Kimbro, 2015).

In the instance of the educational level, those who had achieve high education beyond compulsory years of schooling lead to a lessened probability of struggling to meet their mortgage or rent repayments, but its positive effect was small for those without arrears, albeit it was statistically significant. It seems plausible to expect that higher educational attainment is often correlated with higher income, which can provide a buffer against financial strain. Similar conclusions have been reached by Squires et al. (2016). Likewise, a better financial literacy helps

individuals make more informed decisions regarding housing costs (Lusardi and Mitchell (2014) support this notion). Now, concerning those without arrears may already have some level of financial stability, making the incremental benefit of higher education less pronounced but still statistically significant. This aligns with literature discussing diminishing returns to education on income and well-being (Harris, 2007).

Employment status plays a vital role and is also well-aligned with a broad array of literature on socio-economic factors and housing stability. Somewhat surprisingly, the risk of the outcome falling in each of the 4 comparison groups relative to the risk of the outcome falling in the referent group increases strikingly either being employed or not. More particularly, compared with owners without arrears, struggling owner-occupied households have been shown to exhibit higher rates, while unemployment elevates them even more. Once again, our findings are thus in line with the broader scope of existing literature on the subject. To be more accurate, being employed means a stable income but also comes with financial commitments and social expectations, which can sometimes result in more significant financial commitments, like bigger houses or apartments (Gourinchas and Parker, 2002). Besides this, many of these borrowers have suffered a sudden decline in their income, precarious employment conditions, such as moving from a permanent contract to temporary or part-time work, or have a history of unemployment (McCarthy, 2014).

Furthermore, as the number of individuals in a household increase, a negative effect on the probability of being tenant (with or without arrears), is indicated compared to the probability of not having housing cost while the same went for the impact on the probability of belonging to the first category of owners who not being in difficulty with their mortgage or rent payments. One plausible explanation for this behavior is that more people in a household often equate to

more potential earners, which could contribute to a more stable financial environment and reduce the risk of housing cost burdens. Also, it should be referred that stronger effects for tenants are noted, as well as larger households prioritize the stability of homeownership over the flexibility of renting.

Aiming to extract reliable inference, we created a new variable that categorizes the disposable household income by 10 quantiles for each year group, we can observe that lower income associated with quite less risk of belonging to those who are not experiencing arrears on rent payments, since the budget constraint inhibits the low-income families from accumulating the savings necessary for the repayment of a loan or rent. In contrast, higher income at first two categories seems to approach the probability of not having housing cost. Now, looking at the remaining columns of Table 3 which contain the Category [3] and Category [4], the successively smaller coefficients as going to high-income earners could suggest diminishing returns in the profitability of being owners or tenants with arrears in comparison to the probability of being owner without housing cost, but still remains higher. To rephrase it, higher income provides a cushion, allowing these households to either fully own their property or keep up with rent, thereby reducing the probability of incurring housing costs. Observed that, the income effect is noticeable greater in tenants who fall into arrears.

Apart from income background, rooms' number effect was a very important determinant. The number of rooms increases the probability of having housing cost for owners (with or without arrears) and reduces it for tenants (with or without arrears). Owners may opt for larger apartments because their income allows it, and such, housing costs might be less of a financial burden for them. Nonetheless, as the number of rooms increases is more likely to be an owner without housing cost than a tenant. Concerning geographical insights, we also examine how region of residence can have varying impacts on housing costs and financial stability. First

of all, living in Crete and the Islands reduces the likelihood of being an owner with housing cost and increases the likelihood of being a tenant, based on the relative risk ratios. Case in point, Crete and many of the islands have economies heavily reliant on tourism. This can make property ownership more expensive and might lead to more rentals. However, the coefficients that corresponds to both categories of tenants (with and without arrears) are not statistically significant.

The results, also, indicate that for those living in Northern Greece is significantly increased the likelihood of being an owner, whether with or without arrears, as compared to not having any housing cost. Simultaneously, it reduces the likelihood of being a tenant, either with or without arrears. As it was mentioned in the previous section of Data, ownership is highly valued in this region, leading to a higher percentage of homeowners. The same went for the impact on the probability of being tenants in Central Greece, whereas the likelihood of being an owner without arrears is reduced. Apart from regional dummies relative to the place of residence, we also to display the differences from the division into rural and urban residents. So, as demonstrated by the analysis, people in rural areas are much less likely to be either tenants or owners with arrears. This could possibly be due to fewer available financial services, lower income levels, or cultural factors that de-emphasize debt. Likewise, the pseudo R^2 equals 0.023, meaning that the variance of the dependent variable that can be explained from the independent ones is almost 2,3%.

Table 3. *Relative risk ratios on each type of constraint household*

Variable	[1] Owners-No Arrears	[2] Tenants-No Arrears	[3] Owners-Arrears	[4] Tenants- Arrears
Age	.961 *** (.001)	.926 *** (.000)	.962 *** (.001)	.933 *** (.001)
Female	1.113 ** (.053)	1.007 (.031)	1.199 *** (.077)	1.015 (.043)
Married	1.262 *** (.038)	1.071 *** (.028)	1.220 *** (.051)	1.013 (.037)
University	1.087 *** (.035)	1.033 (.028)	.684 *** (.035)	.614 *** (.026)
Employed	1.314 *** (.055)	1.077 ** (.033)	2.077 *** (.122)	1.702 *** (.073)
Unemployed	1.333 *** (.085)	.907 ** (.041)	3.455 *** (.239)	2.221 *** (.114)
Household Size	.858 *** (.011)	.556 *** (.007)	1.060 *** (.017)	.833 *** (.013)
Income Quantiles				
1	.174 *** (.012)	.622 *** (.031)	1.734 *** (.137)	4.739 *** (.417)
2	.188 *** (.013)	.622 *** (.031)	1.587 *** (.129)	5.014 *** (.439)
3	.247 *** (.016)	.788 *** (.037)	1.252 *** (.134)	1.726 *** (.136)
4	.297 *** (.017)	.806 *** (.037)	1.562 *** (.121)	4.104 *** (.355)
5	.385 (.019)	.860 *** (.038)	1.731 *** (.126)	3.676 *** (.317)
6	.493 *** (.022)	.939 (.040)	1.899 *** (.132)	3.258 *** (.282)
7	.571 *** (.023)	.892 *** (.038)	1.686 *** (.115)	2.567 *** (.224)
8	.681 *** (.025)	.973 (.040)	1.673 *** (.111)	2.136 *** (.189)
9	.845 *** (.028)	.943 (.038)	1.248 (.084)	1.402 *** (.131)
Number of rooms	1.127 *** (.014)	.495 *** (.006)	1.172 *** (.020)	.452 *** (.008)
Crete and Islands	.840 *** (.031)	1.023 (.032)	.839 ** (.046)	1.020 (.043)
Northern Greece	1.184 *** (.033)	.922 *** (.023)	1.238 *** (.052)	.676 *** (.024)
Central Greece	.781 *** (.026)	.913 *** (.025)	1.028 (.047)	.753 *** (.029)
Semi-urban	1.004 (.028)	.872 *** (.020)	.975 (.039)	.806 *** (.027)
Rural	.681 *** (.020)	.352 *** (.009)	.694 *** (.028)	.342 *** (.012)
Pseudo R-squared	.206	.206	.206	.206
Observations	160,347	160,347	160,347	160,347

Source: Income and Living Conditions (EU-SILC, 2007-2022), Hellenic Statistical Authority (EL.STAT.), Public Use Files (Microdata)
Notes: Time-specific regressor was included in the model specification, but they are not presented in the results. The analysis conducted taking into account the average magnitudes of the independent variables; Coefficients estimates using a multinomial logit model to calculate the relative risk ratios of the explanatory variables on the probability of each outcome; Robust standard errors in parentheses, *** p-value<0.01, ** p-value<0.05, * p-value<0.1 denote significance at the 1%, 5% and 10% levels, respectively.

5.1.2 Determinants of housing cost and hardship repayments

Logistic regression of the Model 2 yields the following results in Table 4. As we can see, this table showcases a nuanced understanding of how various socio-economic variables influence the probability of a household incurring housing costs in Greece. Starting with the age, we observe that it has a minor but statistically significant negative effect on the likelihood of having housing costs, suggesting that as people age, the probability of incurring these costs slightly decreases. Older individuals, having more experience and exposure to financial planning, are often more adept at managing household finances, including minimizing housing costs (Henager and Cude, 2016). Also, the result of reducing the likelihood of holding housing cost with age echoes the findings of Yilmazer and DeVaney (2005). The estimated coefficient for the females in model 2 is also displayed in second row of Table 4, but it doesn't have a statistically significant impact, suggesting that the gender of the owner or renter, in this case, might not play a crucial role in determining housing costs.

Households led by married individuals are marginally more likely to incur housing costs than their unmarried counterparts. That is extracted by the coefficient of the dummy variable Married, which is statistically significant at 10% significance level. Married couples often plan for long-term stability, which might include investing in a more substantial or better-located dwelling, which could come at a higher cost. This statement is compatible with the results of Parrotta and Johnson (1998). In addition, following the estimations below, the educational level was a statistically significant determinant since indicated a diminishing effect on those with a university education, and as a result, they are less likely to have housing costs. This could perhaps be attributed to higher income potential, which can provide a buffer, allowing for better financial planning and possibly even outright

home ownership, thereby reducing ongoing housing costs among the educated population, confirming Martin (2007). Also, people with university education are statistically more likely to be homeowners rather than renters, and once a mortgage is paid off, the owner incurs fewer housing costs.

When it comes to employment status, the employed and especially the unemployed are substantially more likely to face housing costs. For those who participated in the labor force, makes it easier to qualify for housing loans, so this could eventually fortify the probability of having housing cost. Adding to that, employment often necessitates living in or near urban centers where job opportunities are more abundant, but housing costs are also generally higher. Finally, we can say that the results of the regression confirm that joblessness brings about fewer choices in terms of housing and may be forced to accept less favorable terms, like higher interest rates on a mortgage. Meanwhile, based on our findings, household size reduces the likelihood of incurring housing costs. This could be either due to the possibility that larger families might own their homes, thus incurring fewer recurring housing expenses or due to economies of scale that larger households are often benefited from. In the second case, the cost per person for housing decreases as the household size increases. So does the number of rooms in a dwelling.

A deeper dive into the income quantiles reveals a consistently negative association between household income levels and housing costs. Another noticeable pattern is the substantial reduction in the income effect on the probability of having housing, as indicated by the coefficients of quantiles, as income goes up, whereas the smallest income class experiencing the highest negative marginal effect compared to the wealthiest households, underscoring the financial challenges faced by lower-income households concerning housing. The difference in the probabilities of having housing cost due to the area of residence is demonstrated by the dissimilar marginal effects of the regions. Firstly, the coefficient of Crete and

Islands and Central Greece are statistically most significant. As we can see, households in Crete and the Islands and Central Greece, and rural areas are notably less likely to have housing costs in comparison to households in Attika. This might reflect the lower cost of living or different housing patterns in these areas. Now, estimating the variations in household financial decisions between urban and rural areas, it should be referred that rural areas are notably less likely to take over a loan or rental in contrast with urban areas. Rural dwellers are more likely to be homeowners who might be engaged in some form of agriculture or livestock rearing, which could offset some housing costs. Also, income discrepancies between urban and rural samples explain the above statement. The overall model seems to account for 21.5% of the variability in the outcome, as indicated by the Pseudo R-squared value. The sample size for this study is extensive, with over 160,000 observations.

Table 4. *Average marginal effects on the likelihood of having housing cost*

Variable	
Age	-.008 *** (.000)
Female	-.002 (.003)
Married	.004 * (.002)
University	-.007 *** (.003)
Employed	.030 *** (.003)
Unemployed	.049 *** (.004)
Household Size	-.044 *** (.001)
Income Quantiles	
1	-.069 *** (.004)
2	-.054 *** (.004)
3	-.035 *** (.004)
4	-.042 *** (.004)
5	-.034 *** (.004)
6	-.023 *** (.004)
7	-.028 *** (.004)
8	-.018 *** (.004)
9	-.017 *** (.004)
Number of rooms	-.044 *** (.001)
Crete and Islands	-.007 ** (.003)
Northern Greece	-.000 (.002)
Central Greece	-.020 *** (.002)
Semi-urban	-.015 (.002)
Rural	-.103 *** (.002)
Pseudo R-squared	.215
Observations	160,347

Source: Income and Living Conditions (EU-SILC, 2007-2022), Hellenic Statistical Authority (EL.STAT.), Public Use Files (Microdata)

Notes: Time-specific regressor was included in the model specification, but they are not presented in the results. The analysis conducted taking into account the average magnitudes of the independent variables; The dependent variable, index0, takes the value 1 if a household has housing cost and the value 0 differently; Coefficients estimates using a logit model; Robust standard errors in parentheses, *** p-value<0.01, ** p-value<0.05, * p-value<0.1 denote significance at the 1%, 5% and 10% levels, respectively.

Table 5 presents the estimation results of the Model 3 described in the previous

section, that provide an overview of how specific variables are associated with the likelihood of falling into arrears with loan or rent repayments. As the individual get older, it is slightly more likely to experiencing hardship with loan or rent repayments. A version of this statement can be that older individuals may be on a fixed income, such as pensions, which may not keep up with either rising living expenses or increasing healthcare costs that can eat into their budget, making it more difficult to make loan or rent payments. However, the analysis of Jian Xiao and Yao (2014) investigated the exact opposite result. According to them, younger households facing more financial distress than older ones. With respect to gender, the second row uncovers a non-statistically significant effect on the probability of dealing with mortgage or rent payment difficulties.

Following the results bellow in Table 5, being married has a noticeable dampening influence in the likelihood of struggling to keep up payments. However, this decrease is not statistically significant. Conversely, previous work supports that marital relationship strengthens the effect of financial knowledge on financial well-being, as well as both partners typically contribute to financial decision-making and can diversify financial risks better. So, when both are financially knowledgeable, they can complement each other's strengths and weaknesses, resulting in better decisions (Iramani and Lutfi, 2021). Having a highest achievement than the compulsory education is associated with a substantial decrease in the likelihood of being a struggling household. This effect is statistically significant at the 1% level. Individuals with more than a basic level of education are more likely to secure jobs with better pay and inculcate a culture of planning for the future and being aware of the consequences of financial decisions, thereby providing a buffer against financial struggles. So, financial literacy includes understanding the importance of tracking income and expenses, and how it affects their ability to secure loans or rentals, managing them effectively (Singh, 2014).

Also, the employment status had a considerable impact on the examined probability. Looking at the estimates of the employment, we detect that being employed is associated with a significant decrease in the likelihood of experiencing hardship. Something noteworthy is that not all employment is equal. Low-wage jobs, gig work, or part-time employment may not provide enough income to meet financial commitments, leading to financial hardship (Kalleberg, 2011). Also, employment might allow access to credit, which if mismanaged, can lead to a cycle of debt. In terms of unemployment, a loss or significant reduction in income, making it difficult to meet their mortgage or rent repayments. This is also confirmed by the positive relationship shown by the coefficient obtained for the specific variable, which means that unemployment magnifies the likelihood of being in arrears of mortgage or rent repayments. So does the household size, as well as for each additional member in a household, the likelihood of experiencing hardship increases. This increase is statistically significant at the 1% level.

The coefficients under income quantiles represent the effect of different income levels on the likelihood of experiencing hardship. Different coefficients for different income quantiles mean that the impact of financial hardship is not uniform across income levels. For instance, for individuals in the lowest income group (1st quantile), a much higher increase in the mortgage or rent burdens is observed in contrast with the 10th quantile, which includes the highest income class. The effects decrease as income quantiles increase. So, as higher one's income is, the less one is leading to hardships. This is consistent with broader literature in economics and sociology that highlights the financial stress that housing can put on lower-income families, which their income is insufficient to meet the housing costs. Our results, also, demonstrate that each additional room in a household is associated with a small increase in the likelihood of experiencing hardship, which is statistically significant at the 10% level.

The coefficients under geographic regions show the impact of living in specific geographic regions on the likelihood of experiencing hardship. For example, living in Northern Greece is associated with a significant decrease in the likelihood of having struggled to afford housing cost. The absence of a statistically significant effect for other regions suggests that, within the scope of our study, geography may not be a universal determinant of hardship. Lastly, we note that the influence of the semi-urban population continues to be insignificant, whereas the estimation results in the row of Rural variable reveal a positive and significant at the 1% significance level coefficient of the interaction between the area and the likelihood of experiencing hardship with loan or rent repayments. More particularly, for those who live in rural areas, it is more probable to fall into arrears in comparison to urban area living. From the information concerning the R^2 that can be extracted from the above table (Table 5), we can say that 10,6% of the variation in the dependent variable, namely the index1, is explained by the independent variables of the logistic regression model.

Table 5. *Average marginal effects on the likelihood of experiencing hardship with loan or rent repayments*

Variable	
Age	.001 *** (.000)
Female	.003 (.007)
Married	-.009 (.006)
University	-.096 *** (.006)
Employed	.085 *** (.007)
Unemployed	.181 *** (.009)
Income Quantiles	
1	.417 *** (.010)
2	.354 *** (.010)
3	.321 *** (.010)
4	.252 *** (.009)
5	.229 *** (.009)
6	.192 *** (.008)
7	.156 *** (.008)
8	.115 *** (.007)
9	.048 *** (.006)
Number of rooms	.005 * (.002)
Crete and Islands	-.003 (.007)
Northern Greece	-.028 *** (.006)
Central Greece	-.002 (.006)
Semi-urban	-.006 (.005)
Rural	.022 *** (.006)
Pseudo R-squared	.106
Observations	39,049

Source: Income and Living Conditions (EU-SILC, 2007-2022), Hellenic Statistical Authority (EL.STAT.), Public Use Files (Microdata)

Notes: Time-specific regressor was included in the model specification, but they are not presented in the results. The analysis was conducted taking into account the average magnitudes of the independent variables; The dependent variable is the index1 which takes the value 1 if a household has experienced hardship with loan or rent repayments and the value 0 differently; Coefficients estimates using a logit model; Robust standard errors in parentheses, *** p-value<0.01, ** p-value<0.05, * p-value<0.1 denote significance at the 1%, 5% and 10% levels, respectively.

5.1.3 Determinants of default in housing

Next, what is going to be analyzed is a model of households' mortgage or rent default incorporating disposable income and the remaining independent variables. This analysis is considering a panel data type. Apart from the examination of all households, I also focus on those who are taking over a loan or having rent repayments. So, concerning the complete set of variables, we estimate two same models. The first one was estimated on the full sample (Column [1]), while the second one was estimated on a sample where households taking over either mortgage or rental (Column [2]), with some significant conclusions obtained, as well as quantifies the effects of mortgage payers' and tenants' characteristics on default.

As Table 6 delves into the likelihood of Greek households defaulting on their housing-related obligations, in terms of the general population, we can observe that the greater the age, the lower the likelihood of default. However, for households with housing costs, the increase in age is positively correlated with the default probability. This could be described as an 'inverse U-shaped relationship,' where age initially decreases the likelihood of default but then increases it after peaking at a certain age, particularly for those with existing housing costs. These diverging patterns reflects the evolving financial risks people face they go through different life stages. On one side older individuals may have more stable income sources and have finished paying off their mortgages, and as a result, negatively, associated with the propensity to default, but on the other side, for those with ongoing housing costs, the increase in age has contributed in to increased default rates due to retirement and a subsequent reduction in income. So, early on, as individuals take on more responsibilities, such as expanding their families, they tend to acquire larger debts, such as a mortgage for a bigger home, which increases their financial burden. However, this risk starts to diminish later in life as these debts get paid down and family-related expenditures decrease, often after children

become financially independent (Alfaro and Gallardo, 2012).

Undoubtedly, as we can see, gender did not indicate a significant effect on the default probability in both overall and selected sample with the implementation of OLS regression models. In the general population, being married increases the likelihood of default, which is counterintuitive to the often-held belief that dual-income households (which are common in marriages) would have a lower risk of financial default. This suggests that the added financial responsibilities or pressures associated with marriage might outweigh the benefits of a potentially larger combined income (Çığşar et al., 2019). Now, among households that have housing-related expenses, being marital status still has a positive association with the likelihood of defaulting, but this association is not statistically significant. So, for households already burdened with housing costs, being married does not significantly alter the risk of default. Turning to university versus non-university educated individuals, having a university degree significantly reduces the likelihood of financial default for individuals in both sample groups, which aligns with the broader understanding that higher education often leads to better job opportunities and financial stability. Something very interesting that is worth to be mentioned, is that the risk-reducing effect of university attainment is almost three times as potent for households that are burdened with housing expenses. This might imply that the skills, networks, or opportunities that come with higher education are especially valuable for navigating the financial complexities that come with housing costs, confirming once again the study of Alfaro and Gallardo (2012).

Both employed and unemployed categories indicate a higher probability of default in both groups, with the effects being particularly pronounced for the unemployed, especially in the households with housing costs group. Instead, having a job is not necessarily a guarantee of financial stability. It may imply that wages are not sufficient to cover housing costs, or that the employed group may have

other financial commitments impacting their ability make regular mortgage or rent payments. Unsurprisingly, unemployment is associated with a higher risk of default. Lack of a regular income makes it challenging to meet recurring financial obligations like mortgage or rent. As we mentioned, for the comparability of the results, the risk seems even more significant for the unemployed in this subset. Turning to the whole sample, this could be due to a lack of financial buffers like savings, or other forms of social security benefits that could otherwise alleviate housing costs, while the subset with housing costs may naturally be more financially stretched, thereby, being closer to the financial edge, and more sensitive to disruptions like job loss.

There's no significant effect on the overall population, but for households with housing costs, an increase in household size could mean more mouths to feed and higher utility costs, among other expenses. So, these additional costs could make it more difficult to meet housing costs, thereby increasing the likelihood of default (Jian Xiao and Yao, 2014; Moulton et al., 2015). A clear understanding of the relation between the default probability and income is demonstrated, while as income increases, results a noticeable dampening influence on the effect on probability for both sample groups. Notably, we find pronounced increases in the predicted marginal effects of the regression runs under the condition that households have housing cost. Numerically speaking, a massive 45.8% increase in the likelihood of default for the lowest-income families with housing cost reflects severe financial constraints. All coefficients, regarding the income quantiles, are statistically significant at level 1%. Moreover, the number of rooms in a dwelling seems to have contrasting effects. In the broader population, having more rooms could be a proxy for wealth or stable income, which plays an important role in deterring default. More rooms could also mean the potential to rent out space for additional income, providing a safety net against financial hardship. And, for the comparability of the results, having more rooms in households with housing

cost could signify higher housing costs (e.g., higher mortgage rates or rents, more significant property maintenance, higher utility bills) that might not be offset by income, thereby resulting in default in payment of housing loan instalments or rental.

With regard to regional dummies, all regions relative to the location included the full sample are illustrated in the Table 6 tended to adversely affect the default risks in comparison to Attika. This suggests that being in a region other than Attika generally appears to a protective effect against default, but this lower risk of default is not evident when looking only at households that already have housing costs, due to the lack of statistical significance. So, the regional location does not have a substantial impact on those who are already financially committed, possibly because other factors (like income, age, employment status) become more critical determinants of default risk for this group. Lastly, in the overall sample, living in rural areas appears to be a protective factor against default. This could be due to a variety of reasons such as lower cost of living, community support, or different lifestyle factors that make it easier to manage expenses. However, contrary to the general population, households with housing costs in rural areas seem to be at higher risk of default. This could be due to a lack of access to high-paying jobs, lower economic activity, or higher transportation costs that could stretchUdoh (2008). The penultimate row of the table reports that the model specifications in columns [1] and [2] explain accordingly 11.2% and 19.5% of the variability in default probability. The analysis of all households is based on a relatively large sample of 19,925 observations, whereas the subset after the selection procedure of households with housing costs is about a quarter of the full sample. So, table 6 below illustrates the previously described results.

Table 6. *Probabilities of default estimates*

	All Households	Households with cost
Variable	[1]	[2]
Age	-.003 *** (.000)	.001 *** (.000)
Female	.004 (.006)	.006 (.017)
Married	.018 *** (.004)	.011 (.013)
University	-.027 *** (.005)	-.074 *** (.014)
Employed	.033 *** (.006)	.110 *** (.016)
Unemployed	.096 *** (.009)	.195 *** (.022)
Household Size	.000 (.002)	.057 *** (.005)
Income Quantiles		
1	.105 *** (.008)	.458 *** (.026)
2	.104 *** (.008)	.403 *** (.024)
3	.090 *** (.008)	.351 *** (.024)
4	.090 *** (.008)	.336 *** (.023)
5	.078 *** (.007)	.275 *** (.022)
6	.064 *** (.007)	.241 *** (.021)
7	.057 *** (.007)	.199 *** (.020)
8	.048 *** (.007)	.147 *** (.019)
9	.013 *** (.007)	.068 *** (.019)
Number of rooms	-.009 * (.001)	017 *** (.006)
Crete and Islands	-.008 * (.005)	.002 (.016)
Northern Greece	-.028 * (.005)	-.021 * (.013)
Central Greece	-.009 * (.005)	.000 (.014)
Semi-urban	-.009 ** (.004)	-.008 (.012)
Rural	-.033 *** (.004)	.043 *** (.013)
Pseudo R-squared	.112	.195
Observations	19,925	5,051

Source: Income and Living Conditions (EU-SILC, 2007-2022), Hellenic Statistical Authority (EL.STAT.), Public Use Files (Microdata)

Notes: Time-specific regressor was included in the model specification, but they are not presented in the results. The analysis conducted taking into account the average magnitudes of the independent variables; The dependent variable is the probability of default; Coefficients estimates using the OLS method; Robust standard errors in parentheses, *** p-value<0.01, ** p-value<0.05, * p-value<0.1 denote significance at the 1%, 5% and 10% levels, respectively.

Chapter 6

Conclusions

In order to conduct our research and make our deductions based on the research question, we were based on evaluation of 160,348 housing cases in Greece, combining data on EU-SILC with detailed borrower and household characteristics. Initially, we identify borrowers and tenants who face debt and rent paying challenges, respectively. By distinguishing between debt-related challenges for borrowers and rent-related challenges for tenants, we can draw more specific conclusions about the unique factors affecting each group. Firstly, we estimate the probability of a household belonging separately in each category of constraint households (with or without arrears) using a multinomial model and secondly conduct an analysis of the relationship between the likelihood of having housing cost and the likelihood of struggling to keep up their mortgage or rent repayments, using two logit models, respectively. Afterwards, within a balanced panel data set, we've chosen to implement a regression using the Ordinary Least Squares method. This approach allows to uncover key factors that significantly influence default probabilities. The analysis depends on numerous individual and household-level features. Among those that matter is age, university attainment, household size, income quantiles and urban-rural areas.

Designing policies to address a mortgage arrears crisis is a complex task that

requires a comprehensive understanding of the problem from multiple angles. A dynamic micro-simulation to simulate the likely future repayment profiles and assesses the proportion of total debt repaid could be a useful tool for policy makers. Success lies in tailor-making loans in tune with the risk profiles of individual customers. So, government and private lenders to assess the risk of default and develop appropriate financial management strategies to minimise risk. An important turning point of the presented research is that our findings are likely to provide and modulate policy interventions concerning housing debt. So, based on the empirical findings, we extracted some interesting conclusions.

The study established that there existed a negative correlation between the age and the likelihood of having housing cost. The results also show that many borrowers or tenants experiencing arrears are currently employed. This shows that the current situation, and efforts to prevent a further deterioration, requires more than simply targeting unemployment. Rather, such efforts should also aim to strengthen overall labour market conditions, like job security or emergency housing funds, that can be mobilized to help households in situations of financial hardships. Providing financial counseling and planning resources to at-risk populations can also be a proactive measure. Moreover, after examining our data, we come to the conclusion, that disposable income plays a significant role in the likelihood of a household having housing cost and stands out as the variable with the highest incidence. The results indicate that, within the whole and selected sample of households with cost, very low or lower ends of income is associated with higher default. This is a comprehensive set of policy implications derived from the study.

References

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., and Evanoff, D. D. (2011). Financial counseling, financial literacy, and household decision-making. *Financial literacy: Implications for retirement security and the financial marketplace*, 181.

Alfaro, R. and Gallardo, N. (2012). The determinants of household debt default. *Economic Analysis Review*, 27(1):27–54.

Blau, F. D. and Kahn, L. M. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of economic literature*, 55(3):789–865.

Browning, M. and Lusardi, A. (1996). Household saving: Micro theories and micro facts. *Journal of Economic literature*, 34(4):1797–1855.

Bucks, B. and Pence, K. (2008). Do borrowers know their mortgage terms? *Journal of urban Economics*, 64(2):218–233.

Campbell, J. Y., Giglio, S., and Pathak, P. (2011). Forced sales and house prices. *American Economic Review*, 101(5):2108–2131.

Çığşar, B., Ünal, D., et al. (2019). Comparison of data mining classification algorithms determining the default risk. *Scientific Programming*, 2019.

Desmond, M. and Kimbro, R. T. (2015). Eviction's fallout: housing, hardship, and health. *Social forces*, 94(1):295–324.

Gerardi, K., Herkenhoff, K., Ohanian, L. E., and Willen, P. (2013). Unemployment, negative equity, and strategic default. *Available at SSRN 2293152*.

Gourinchas, P.-O. and Parker, J. A. (2002). Consumption over the life cycle. *Econometrica*, 70(1):47–89.

Harris, D. N. (2007). Diminishing marginal returns and the production of education: An international analysis. *Education Economics*, 15(1):31–53.

Henager, R. and Cude, B. J. (2016). Financial literacy and long-and short-term financial behavior in different age groups. *Journal of Financial Counseling and Planning*, 27(1):3–19.

Iramani, R. and Lutfi, L. (2021). An integrated model of financial well-being: The role of financial behavior. *Accounting*, 7(3):691–700.

Jian Xiao, J. and Yao, R. (2014). Consumer debt delinquency by family lifecycle categories. *International Journal of Bank Marketing*, 32(1):43–59.

Kalleberg, A. L. (2011). *Good jobs, bad jobs: The rise of polarized and precarious employment systems in the United States, 1970s-2000s*. Russell Sage Foundation.

Lusardi, A. and Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. *American Economic Journal: Journal of Economic Literature*, 52(1):5–44.

Martin, M. (2007). A literature review on the effectiveness of financial education.

McCarthy, Y. (2014). Disentangling the mortgage arrears crisis: The role of the labour market, income volatility and housing equity. *Income Volatility and Housing Equity (January 31, 2014)*.

Mian, A. and Sufi, A. (2009). The consequences of mortgage credit expansion: Evidence from the us mortgage default crisis. *The Quarterly journal of economics*, 124(4):1449–1496.

Modigliani, F. and Brumberg, R. (1954). Utility analysis and the consumption function: An interpretation of cross-section data. *Franco Modigliani*, 1(1):388–436.

Moulton, S., Collins, J. M., Loibl, C., and Samek, A. (2015). Effects of monitoring on mortgage delinquency: Evidence from a randomized field study. *Journal of Policy Analysis and Management*, 34(1):184–207.

Parrotta, J. L. and Johnson, P. J. (1998). The impact of financial attitudes and knowledge on financial management and satisfaction of recently married individuals. *Journal of Financial Counseling and Planning*, 9(2):59.

Singh, U. (2014). Financial literacy and financial stability are two aspects of efficient economy. *Journal of Finance, Accounting and management*, 5(2):59.

Squires, A., Uyei, S. J., Beltrán-Sánchez, H., and Jones, S. A. (2016). Examining the influence of country-level and health system factors on nursing and physician personnel production. *Human resources for health*, 14(1):1–10.

Udoh, E. J. (2008). Estimation of loan default among beneficiaries of a state government owned agricultural loan scheme, nigeria. *Journal of Central European Agriculture*, 9(2):343–352.

Yilmazer, T. and DeVaney, S. A. (2005). Household debt over the life cycle. *FINANCIAL SERVICES REVIEW-GREENWICH-*, 14(4):285.