TALLINN UNIVERSITY OF TECHNOLOGY
School of Business and Governance
Department of Economics and Finance

Kadri-Ann Freiberg

DETERMINANTS OF HOMEOWNERSHIP RATE IN ESTONIA

Bachelor's thesis
Programme Applied Economics, specialisation Economic Analysis

Supervisor: Simona Ferraro, PhD

Tallinn 2020

I declare that I have compiled the paper independently
and all works, important standpoints and data by other authors
have been properly referenced and the same paper
has not been previously presented for grading.
The document length is 11 885 words from the introduction to the end of conclusion.

Kadri-Ann Freiberg……

(signature, date)

Student code: 180627TAAB

Student e-mail address: kafreiberg@gmail.com

Supervisor: Simona Ferraro, PhD

The paper conforms to requirements in force

……

(signature, date)

Chairman of the Defence Committee:

Permitted to the defence

……

(name, signature, date)

TABLE OF CONTENTS

ABSTRACT

Several studies have discussed the determinants of homeownership rates in the past, yet the majority of research has focused on cross-country analysis rather than on a specific region. The aim of the thesis was to explore the determinants of homeownership rates in Estonia using time series data for the period from 2006 to 2019. Regression analysis was performed in Gretl software using the method of ordinary least squares. The dependent variable was homeownership rates and the independent variables were income, mortgage interest rates, house price index and price to rent ratio.

A statistically significant model was estimated with explanatory power of 98%. Homeownership rates in Estonia were found to be positively related to mortgage interest rates and negatively related to income and house prices. The relative price of owning i.e. price to rent ratio was not statistically significant. Implications of the estimated model are discussed.

Models for homeownership rates of different household types and income groups are estimated and used for robustness check. Analysis by subgroups resulted in statistically insignificant models with one exception, which was the model for lower income households. Limitations and implications and options for further research are discussed.

INTRODUCTION

After the collapse of the Soviet Union and the re-establishment of independence in 1991, which brought upon the privatisation of land and real estate, the Estonian housing market has experienced a rapid growth in prices, accompanied by substantial economic development. While the homeownership rate in the country has increased in line with the overall trend in other Organization for Economic-Cooperation and Development (OECD) countries, the proportion of homeowners has fallen after the 2008 economic crisis and has yet to reach the pre-crisis levels (Eurostat, table *Distribution of*…).

Past literature and empirical evidence has linked homeownership to an array of socio-economic benefits, both at a micro level for households and in the society as a whole. These include increases in well-being and wealth (e.g. Coulson 2002; Gatzlaff *et al.* 1998; Haurin 2001; Rohe and Stewart 1996), educational outcomes (Green and White 1997), life satisfaction and psychological health (Evans *et al.* 2003; Rohe, Lindblad 2013), reduced crime rates (Lauridsen *et al.* 2013; Dietz, Haurin 2003), stability (Coulson 2002; Dietz, Haurin 2003; Rohe, Stewart 1996) and a more active and involved citizenry (Di Pasquale, Glaeser 1999).

Public policy in developed countries, including Estonia, is generally geared towards facilitating homeownership, whether through preferential tax treatment (in Estonia, it is possible to deduct mortgage interest payments from taxable income) or by alleviating credit constraints (e.g. KredEx). As mentioned above, real estate offers households an opportunity to invest and accumulate wealth, enhancing financial stability and improving households' economic outlook.

As real estate has become one of the primary class of assets for many households in Estonia, the thesis aims to provide insight into Estonian housing market utilizing the most recent data available. The topic under study is relevant because due to the benefits listed above, increasing homeownership rates is generally a favorable and desirable policy (Coulson 2002, Andrews, Sánchez 2011). Thus, it is important to understand the factors affecting tenure choice in order to

evaluate socio-economic and demographic determinants and public policy tools that may enhance owner-occupancy.

Although the availability of literature and data related to housing opportunities and tenure status on an international level has increased over time, on a local level there is a lack of research related to the subject. Moreover, an overwhelming amount of previous studies focus on international cross-sectional comparisons between countries, as opposed to using time series data of a specific region. At this point, no studies have been undertaken for Estonia which provided a possible research opportunity to be developed in the current thesis.

The purpose of this thesis is to explore the determinants of homeownership rate in Estonia through econometric modelling. The objective is to determine, if and to what extent, the homeownership rate in Estonia from the year 2006 to 2019 is influenced by mortgage market conditions, socio-economic and demographic factors.

The aim of the thesis is to answer the following questions:

1) According to previous studies, what are the most common factors that have an impact on homeownership rates?
2) Is there a statistically significant relationship between homeownership and one or more of the studied independent variables (income, mortgage interest rates, house price index, price to rent ratio)?

Four hypotheses are proposed:
H1: Mortgage interest rates have a negative relationship with homeownership rates
H2: Household income has a positive relationship with homeownership rates
H3: House prices have a negative relationship with homeownership rates
H4: Price to rent ratio has a negative relationship with homeownership rates

The author's objective is to summarize earlier literature, provide statistics and findings on the aforementioned subject and to use econometric modelling as the main tool for inquiry. In order to answer the research questions, regression analysis is performed, where homeownership rate is set as the dependent variable and the chosen explanatory factors are set as the independent variables. Thereby, it is possible to conclude which, if any, of those factors and to what extent have an impact on homeownership.

Multivariate regression analysis will be performed in econometric package Gretl (Gnu Regression, Econometrics and Time-series Library). The data will mainly be retrieved from Eurostat, OECD and Statistics Estonia (Statistikaamet) that have reliable and up to date information on the chosen variables. As the homeownership rate data in Eurostat is available starting from 2006, this will be the first year in the time series.

The thesis will be divided into three sections. Firstly, in order to provide an answer to the research questions posed, the paper provides a theoretical background for the Estonian housing market, benefits of homeownership, summarising previous studies and relevant factors. Secondly, a research methodology is developed in effort to analyse the determinants of homeownership in Estonia. Thirdly, the research is executed, modelled and results discussed. Finally, the paper is concluded with possible limitations and implications.

1. THEORETICAL BACKGROUND

This chapter aims to give an overview of the real estate market and its development in Estonia in the previous decades as well as the theoretical background of the benefits and consequences of homeownership in society as a whole. The last subsection summarises some of the previous empirical studies exploring the factors affecting ownership rates.

1.1. Real estate market in Estonia

Prior to the collapse of the Soviet Union, housing estates on both sides of the Iron Curtain started being subject to privatisation and a general reduction of state control (Marcuse 1996). Such shifts always tend to echo wider trends in societies and facilitate fundamental changes on a larger scale.

The Estonian private housing market started growing in the wake of Estonian restoration of independence from the former Soviet Union, in August 1991. The dominant role of the state was even more prevalent in Estonia, than in most of the other post-soviet Eastern bloc countries. Around 70% of dwellings belonged to the state or local municipalities and there was no private ownership of land (Purju 1996). Regardless, in Estonia, the transfer of assets from the state to individuals was done in a rather rapid strategy, compared to some of its counterparts like the Czech Republic and Poland (Kährik *et al.* 2004).

In 1993, the "Law on Privatisation of Dwelling Rooms" was created and the voucher (privatisation securities) system introduced. There were two types of privatisation securities that could be used to gain ownership of dwellings - national capital bonds, the so-called "yellow cards" and compensation securities. Every citizen or permanent resident of Estonia who was at least 18 years old could apply for the national capital bond (NCB). The basis for calculating the value of the bond were the working years between 1945 and 1992 and the number of children, each of whom added 5 working years to the parent's bond (Purju 1996). Tenants occupying state owned dwellings were given the option to acquire the housing unit they resided in on very

affordable terms. Using the NCBs to cover the cost of the dwelling, their expenses were limited to the state duty and contract fee that mostly added up to around 1 percent of the total value of the property (Kährik *et al.* 2004). Apartment buildings were transformed into condominiums, where each tenant owned their own unit and a fraction of land under the building and around it. Land was not privatised on the same terms, however, and the transfer of surrounding territory took place later than the transfer of dwellings[1] (Tuvikene 2019).

The other option, compensation securities (CS) were regulated by the "Law on Determination of and Compensation for the Value of Unlawfully Expropriated Property" (adopted by the Parliament in 1993). In 1940, when Estonia was incorporated into the Soviet Union, land and property were nationalised. CS were issued to former owners and their heirs, in cases where it was impossible to return the nationalised property in its physical form, or where they do not want it returned (Purju 1996).

At the beginning of 1994, privately owned dwellings made up 29% of housing stock (the rest was owned by the state or local municipalities). Within 5 years, as a result of the privatisation scheme explained above, that figure had increased to 93% (Tuvikene 2019) and stood at 98% according to the last Population and Housing Census in 2011 (Statistics Estonia, Conventional dwellings..). National capital bonds expired in 1996 and compensation securities in 2005, by which time 16.74 billion kroons worth of bonds and securities had been issued (Ministry of Finance 2005). Through reform, in less than 10 years, a highly regulated housing sector was transformed to an almost 100% privately owned *laissez faire* market system.

Going forward, real estate prices in Estonia reached their peak in 2007, when the OECD house price index surged to 133.4 (2015 = 100). Due to the 2008-2009 worldwide economic crisis, the index dropped to 71.2 in 2009, almost halving the market value of private property and home equity. During the crisis, it became evident that house prices were overly inflated and overvalued. Combined with other effects of the recession like unemployment (in the second quarter of 2010, unemployment rate in Estonia reached 19.8%, compared to an average of 10% in EU-27) (Rosenblad 2011), this brought upon hardship for a substantial amount of homeowners, many of whom had used their dwellings as collateral for mortgage.

Following the recession, as a general trend the economic indicators in the EU countries have constantly improved. The real estate market has continued to boom, salaries and investments

[1] Privatisation of land is out of the scope of this thesis and will not be discussed.

have been increasing. In the EU, property values have gone up by 23% percent since 2009 (Eurostat, House price ...) and owning real estate consequently has become one of the primary class of assets for households in Estonia. Post recession house prices made a substantial recovery in the following years and in Estonia, the house price index currently stands at 112.5 while homeownership rate as of 2019 was 81.7% (Eurostat, distribution of....). Real estate prices have continued to increase up to the current year COVID-19 crisis, which cooled down the market slightly but the real effects are yet unclear. Real estate experts (e.g. Eliste and Habal, cited in Oja 2020) have argued that prices may fall to some extent in the following months, but the effects of the crisis will be rather limited compared to the last recession.

1.2. Homeownership and its benefits

Past literature and empirical studies have linked homeownership to a host of social, economic and psychological advantages and it is believed to bring about positive spillovers for society. In addition to the more obvious ones for the owners themselves, there are also some that are not as straight-forward.

In his 2002 article "Housing Policy and the Social Benefits of Homeownership", Coulson argues that it is impractical and difficult to measure the monetary return owner occupancy provides for the community and challenges whether it justifies the subsidies arising from the U.S tax code, a case can still be made for providing tax incentives for owner occupied housing. That is, if homeownership provides benefits to people other than the homeowners themselves.

Coulson's concise overview of the external benefits that homeowners offer to their neighborhood includes the idea that homeownership in some way could change people's behavior. He explains this phenomenon through the removal of the moral hazard associated with renting[2]. In this case, the moral hazard means behaving "badly" as renters have less incentive to behave "well". Ownership mitigates the aforementioned hazard due to the fact that changing tenure status from renter to owner is often time consuming and a financially significant commitment. These costs are not something households could bear constantly, say annually. Thus, owners tend to stay

[2] Coulson explains: "*Moral hazard often comes up in the analysis of insurance markets, where being insured against bad outcomes (illness, traffic accidents) tempts people into risky behavior that leads to these same outcomes."*

longer in their chosen units and because of that, the payoff from good behavior will be larger and the neighborhood more stable.

This brings us to the stability factor. Due to the high transaction costs relative to renting, homeowners are less mobile than renters. Dietz and Haurin (2003) found that throughout the numerous studies conducted on micro-level consequences of homeownership, this trend holds for different ages, education levels, genders, races and incomes. The fact that homeowners are less likely to move means that they become more embedded and invested in their community (Yun, Evangelou 2016). One of the best ways to stabilize areas in decline and derelict areas is to increase the proportion of owner occupied dwellings. In their census data analysis, Rohe and Stewart (1996) concluded that longer lengths of tenure and less residential mobility indicate greater property value appreciation.

Stability and ownership induces better maintenance of property, which is one of the key differences in the behavior of tenants and owners. Coulson (2002) argues that regardless of the time spent, owners reap the benefits of any improvements and maintenance done on the housing unit. On the contrary, renters rarely see a return on the care put into their rental homes and hence, have less incentive to do so. It is in the owner's best interest to maintain their property and increase the value of the residence when the time comes to sell. Research comparing the differences in price appreciation has found some evidence that owner-occupied housing appreciates more than renter-occupied, primarily due to spending on improvements (Gatzlaff *et al.* 1998).

However, critically thinking, it must be taken into account that the aforementioned maintenance trends do not necessarily always have a causal relationship with homeownership. Homeowners tend to be more often married couples, older and more educated. These traits on their own may bring about better care of the property regardless of tenure status and can be mistakenly associated with the choice to own.

Other social aspects related to homeownership are reduced crime rates (Lauridsen *et al.* 2013; Dietz, Haurin 2003) and a more active and involved citizenry. Di Pasquale and Glaeser (1999) found that, as discussed above, homeownership creates a barrier to mobility and therefore may encourage homeowners to contribute to their community, invest in local amenities and social capital. Homeowners become more informed on local issues and compared to renters, are found to be more involved with community problems. Homeowners are more often members of

non-professional organisations and were also found to be more active voters in local elections, as they have a vested interest in the wellbeing of their neighborhood and local affairs.

Research also shows positive mental health effects for homeowners, which include higher life satisfaction rates (Rohe, Lindblad 2013), sense of control and higher self-esteem (Evans *et al.* 2003). There also seems to be a link between children's success and whether they come from an owner occupied or rental housing. Resident children have been found to achieve higher test scores and have better educational outcomes, which in turn, could lead to higher future income prospects (Haurin *et al*. 2001). Green and White (1997) found that school enrolment and graduation rate is higher for those raised by home-owning parents and offer numerous mechanisms through which these differences can be explained.

First, they argue that purchasing a home requires homeowners to display certain behavioral characteristics and skills. For example, the skills to maintain a home, which include maintenance activities, financial skills to handle mortgage payments and save up for a down payment. Green and White (1997) suggest that these skills pass on to the homeowners' children, who therefore could subsequently become more responsible. Furthermore, the authors also suggest that homeowners have a vested interest to keep the property values and attractiveness of their neighborhood up. Hence, they might apply more disciplinary action and in particular, focus on minimizing socially deviant behavior of their children, as this could impact property values negatively. However, the causal link between homeownership and performance in school is not explicit. The above mentioned characteristics, that children in owners' households demonstrate, could also be the result of residential stability, which in turn raises their educational achievement (Yun, Evangelou 2016). It could also be argued that families who own their home are often wealthier, hence they would have more means to invest in the quality of their offspring's education.

Despite the numerous benefits listed above, a few studies have also discussed the negative impact of policies that promote homeownership. Hilber (2007) points out the very high opportunity costs of these policies in terms of foregone taxes (Follain, Ling 1992; Bourassa, Grigsby 2000, referenced in Hilber 2007, 5), which could counteract the social benefits received. Thus, it would be advisable for policymakers to also apply some kind of cost-benefit approach in assessing if and to what extent a particular policy could potentially assist in homeownership attainment.

1.3. Previous empirical studies on the determinants of homeownership

In recent decades, homeownership has become the subject of extensive economic research. As increasing homeownership is generally a favorable public policy, its determinants have been examined more frequently to provide insight for authorities and policymakers. A comprehensive list of studies exist in the broader literature that have examined factors most commonly affecting ownership rates. Aiming to gather recurring themes, these will be discussed below. In general, past studies (e.g. Struyk, Marshall 1974; Gwin, Ong 2008; Andrews, Sánchez 2011) have mostly been cross-sectional, based on panel data and focused on international differences, as opposed to using time series data of a specific region. With respect to explanatory variables, past tenure choice studies coincide to a large extent. Some form of income measures (individual or aggregate), demographic indicators (e.g. age, marital status), mortgage market conditions (access to credit) and rental market conditions are some of the most common themes discussed when explaining the differences in ownership rates. However, the results have somewhat differed between authors, across regions and time periods.

1.3.1 Income

Income and wealth are arguably some of the most common factors in tenure choice models presented in past literature. As purchasing a home is potentially one of the largest financial transactions and commitments most families will take in their lifetime, for most households this factor tends to surpass other determinants in terms of significance.

Among US studies, some of the earliest ones worth considering are the ones by Carliner (1974) and Struyik and Marshall (1974), both of whom use the 1970 Census Bureau microdata on US households. Those studies are landmarks in early literature as they, albeit being conducted almost 50 years ago, reveal several trends that have been confirmed in later research and still hold relevance today. Whereas Carliner explores different socio-economic determinants that could play a role in homeownership rates, including income, age, marital status, race, family size and location, Struyik and Marshall focus primarily on the relationship between tenure choice and income. While the former author does find income to be positively correlated to ownership rates, the latter digs deeper in analysing the dynamics of these two variables, setting a precedent for future studies.

Struyik and Marshall *a priori* presume that the relationship between income and the probability of homeownership is non-linear and have therefore, in their model, used permanent income in quadratic form. The authors assumed that as the household's income reaches a certain level great enough, the probability of owning asymptotically would approach an upper bound close to unity. As expected, increments of income increase the probability of home attainment but at a decreasing rate at the higher end of the income spectrum. In other words, after a household's income has reached a certain level, each additional monetary unit would have a smaller significance in determining whether the household is an owner-occupant or not.

From microeconomic theory, the income elasticity of demand for becoming a homeowner and diminishing marginal utility were discussed. Struyik and Marshall found families in low and medium income brackets to be more sensitive to changes in income and for incomes above a threshold, the responsiveness started to decrease. They also found that increases in income affect younger families (under 45) more than older families, as even a small boost in income could potentially provide the margin needed for a down payment or to reach a certain credit rating.

Applying some of the same techniques as the former authors, Segal and Sullivan (1998) confirmed Struyik's and Marshall's findings while using a larger sample of the US population census data from 1977 to 1997 (roughly 50 000 households). Their analysis confirmed that even though aggregate homeownership rates rise with real income, they do so at a decreasing rate, having a larger effect on the low end of the income distribution.

Extensive work has been performed by Fisher and Jaffe (2003) in their multivariate cross-sectional study for 106 countries with homeownership rates reported between 1980 and 1999. Their primary sources of data are United Nations and World Bank databases and instead of using microdata on distinct households, Fisher and Jaffe take a different approach using aggregates. They therefore begin with GDP per capita as a general measure of wealth. Controlling for other factors, GDP per capita is found to have a positive yet statistically insignificant relationship to homeownership rates. The square of GDP offers higher explanatory power but its negative coefficient seems to be inconsistent with the general notion set in preceding tenure literature. This also implies that the impact of GDP per capita on homeownership rates is non-linear. The authors infer that all else equal, homeownership rates are likely to be lower at very high levels of income, which may explain the low ownership rates in wealthy nations such as Switzerland or Germany

More recent findings on income are from Andrews and Sánchez (2011), who in their article for the OECD journal analyzed the evolution of aggregate homeownership rates in 15 selected OECD countries from the mid 1990s to mid 2000s. The authors adopted micro economic data decomposition techniques to separate the effects of changes in different household characteristics to the change in aggregate homeownership rates. The impact of changes in real household incomes on homeownership rates did not show the same pattern across countries. In Canada, Denmark and Finland, increments in income raised homeownership rates by 1.5 percentage points while having only minor effects in Germany and other continental European countries.

The latest study on the determinants of homeownership in Europe was published early this year (2020), in which the authors Garcia and Figueira utilized wave 6 (2015) data from the Survey of Health, Ageing and Retirement in Europe (SHARE), containing a sample of 46 003 respondents across Europe. Using the logit model to estimate the probability of an individual being an owner-occupant, several factors were found to contribute to the likelihood of becoming a homeowner (that will be discussed later in this chapter), but in keeping with the spirit of tenure and housing literature, income was found to be positively correlated to the probability of homeownership. Higher level of education had the same positive effect as income, which was previously noted also by Segal and Sullivan (1998), Gyourko and Linneman (1997), Andrews and Sánchez (2011) but will not be discussed separately in this paper as it is a factor highly associated to income and wealth and in some cases, could even be proxied by it.

1.3.2 Demographics

Demographic factors are among other influential elements in explaining homeownership rates. Chambers *et al*. suggest that demographics account for up to 31% of the long-run changes in homeownership rates (Chambers *et al*. 2009)

Age in particular has been heavily stressed in past literature. Carliner (1974) found adjusted homeownership rates more than triple for the youngest age group to the oldest (households with a head under the age of 25 versus families with the head 65 or over). Segal and Sullivan's (1998) results showed a rapid increase in homeownership rates with age up until household heads are around 40 years of age. Confirming historically established viewpoints, the positive relationship between age and homeownership has also been confirmed by Hilber (2007) and Fisher and Jaffe (2003), the latter of whom found the greater the proportion of a country's population in the age group of 15 to 64, the higher the rate of homeownership. During the period studied by Andrews

and Sánchez, on average, population ageing boosted homeownership about 0.75 to 1 percentage points. Their results imply that in the OECD countries studied, all else equal, homeownership rates would have still increased only due to population ageing.

Possible explanations for this trend include the fact that younger households are more often financially constrained, may not have the funds available for a mortgage down payment or may prefer to remain mobile for educational or work related purposes. Garcia and Figueira's study based on SHARE seems to refute this idea of the proportion of homeowners increasing with age. However, the negative relationship determined is possibly explained by the fact that due to data limitations, their sample included households aged 50 and above and is not comparable with the previously mentioned studies.

Besides age, other demographic factors such as marital status, household size and ethnicity have been discussed. For example, Carliner (1974), Hilber (2007), Garcia and Figueira (2020) find ownership rates to be highly correlated with marital status, as married couples tend to have a preference for stability and usually anticipate an increase in household size. In other words, the more frequently a household expects to move, the less feasible it is for them to buy. Carliner's results revealed households headed by married couples had an ownership rate of 71% whereas for those headed by unmarried people the number was only 46%. This could be explained by other factors than just marital status. For instance, households with married heads tend to include children, have higher incomes and with shared finances would have better access to credit.

Likewise to Struyik and Marshall (1974), Carliner (1974), Hilber (2007), Garcia and Figueira find household size to be strongly correlated to the probability to own. In comparison, Fisher and Jaffe find household size to have a positive yet statistically insignificant relationship with ownership rates. For Segal and Sullivan, the results for their studied time period (1977-1997) clearly showed rising homeownership rates among smaller households (without children) while those with more members saw falling rates. This phenomenon could be explained by the fact that younger households, without children, in anticipation of having offspring might purchase homes *prior* to actually extending their families. In terms of household size, Andrews and Sánchez presented varied results among the countries studied having no clear patterns to generalize.

Race and ethnic/immigrant origins have been subject to analysis in a notable amount of studies and is often used as a proxy for socio-economic disadvantage (see Lauridsen *et. al* 2006; Andrews, Sánchez 2011). Carliner (1974) noted a significant difference in ownership rates

between the US's white and African-American households (65% vs 42%). This could be explained by the fact that the African-American households studied by Carliner tended to have smaller incomes and were more often headed by younger adults than white households. The gap in ownership rates between white and non-white households in the US has later been confirmed by Gyourko and Linneman (1997) and Segal and Sullivan (1998). Expanding their research in 1999, Gyourku, Linneman and Wachter c*eteris paribus* find no difference in ownership rates between minorities and whites who have the required financial capacity. Among wealth-constrained households, however, there were substantial differences in the propensity to own. Constrained whites were found to have a 1.7 times greater propensity to own than a household with the same characteristics headed by a minority.

In contrast to studies conducted in the US, Gwin and Ong (2008) utilize aggregate United Nations and World Bank data to provide cross-country analysis of 48 countries between the years 1993 to 1998. In cross country comparison they find race ethnicity to be insignificant in determining homeownership rates. Consistent with previous literature on immigration (see Coulson 1998, Lauridsen *et. al 2006* for a review of determinants of homeownership in Denmark and Andrews, Sánchez 2011 for cross-country comparisons), Gwin and Ong (2008) find immigrants to Europe or North America to be less likely homeowners in their destination countries, regardless of their tenure status in their country of origin.

1.3.3 Mortgage market innovations and public policy

A notable portion of the changes in homeownership rates are not explained by transitions in demographics. Mortgage market innovations and public policy have proven to be another important category in explaining shifts in tenure choice.

Chambers et. al (2009) have accounted for the boom in homeownership rates in the US from 1994 to 2005. Their concise analysis includes mortgage market innovations such as reduction in transaction costs, introduction of new mortgage products and reductions in down payment requirements. The authors found that in the long run, changes in the mortgage market accounted for between 56 and 70% of the increase in homeownership rates during the period studied. A smaller portion of changes was explained with demographics. (Chambers *et. al* 2009)

In the OECD countries studied by Andrews and Sánchez, besides demographics, relaxation of down-payment constraints on mortgages seem to have elevated aggregate homeownership rates

among lower income households, yet rough estimates indicated this change to be comparable with the effect of population ageing (discussed above in chapter 1.3.1).

Consistent with previous studies (see for example Capozza *et al*. 1996; Glaeser, Shapiro 2003), the authors noted that policies such as mortgage interest deductibility tend to be regressive and ineffective in raising homeownership rates as they mostly benefit high income households. Results showed that such tax relief can also inhibit the purchase of a home among lower income households via house price capitalization effect.

At this point, it ought to be mentioned that the discussion above on the varied results on the determinants of homeownership are heavily composed of developed countries. Covering partly the same time period (1993-1998) as Fisher and Jeffe (2003), Gwin and Ong (2008) have focused on homeownership trends in less developed countries. While some of their methods and findings collide with Fisher and Jeffe, Gwin and Ong somewhat expand the scope and improve the methodology of the former paper.

The authors hypothesize that the trends in less developed countries in regards to the choice and opportunity to own do not conform to those in developed states due to institutional, social and political factors. Consistent with previous research their international evidence on developed countries supported the theory that homeownership is sensitive to the price-to-rent ratio as well as the fact that homeownership rates increase in line with increases in income. In addition, Gwin's and Ong's results suggest a negative relationship between interest rates and homeownership in developed countries, as expected.

However, contrary to developed countries that have dominantly been the subject of study in past literature, these relationships did not seem to hold as well in developing countries. What's more, in their sample of countries, Gwin and Ong (2008) intriguingly found low-income countries to have relatively higher homeownership rates than those of higher-income countries. One would assume this phenomenon to be possibly explained by the relatively higher cost of homeownership in high-income countries yet their findings contradicts this hypothesis. Evidence shows that the relative house price-to-rent ratio for developing countries are generally higher than that for developed countries. The authors clearly recognize the need for further research in non-developed countries and suggest the rationale for these differences to be attributed to government assistance programs, subsidies and differences in definitions of formal ownership.

1.3.4 Rental market conditions

Lastly, rental market conditions have been subject to analysis and are associated with the varied homeownership rates in cross-country comparison. As mentioned above by Gwin and Ong (2008), price-to-rent ratio is an important factor one should consider when assessing the cost of becoming an owner-occupant. Other elements related to rental markets include taxation, rent control and availability and composition of housing stock.

The results of Lauridsen *et al.* based on 270 Danish municipalities suggest that short-term changes in house prices and rent control measures have a positive effect on demand for rental units. Medium term-price changes, on the other hand, have a positive effect on demand for homeownership. With respect to other determinants, congestion was found to inhibit higher homeownership rates which relates to the negative effect of urbanisation also discussed by Fisher and Jaffe (2003).

The last study to discuss is that of Hilber (2007) briefly mentioned earlier in this chapter. Contrary to previous research, instead of focusing on socio-economic and demographic variables in determining housing tenure outcomes, Hilber sheds light on other possible location specific characteristics. These include, the impact of the type of accommodation, neighbourhood characteristics (for example, neighbourhood externalities), local housing stock composition on the aggregate level, different tax policies and the importance of public rental housing.

Interestingly, while intergenerational cohesion, proxied by intergenerational cohabitation by Hilber, is often hypothesized to explain the relatively high ownership rates in Southern Europe, in this case it had no discernible effect at all. Of greater relevance are the non-household specific factors. For example, at the household level in the European countries studied, accommodation type turned out to be by far the most important factor in deciding tenure outcomes. Correspondingly, at the regional level the factor of greatest importance was the composition of the housing stock. Hilber suggests the composition of housing stock to partly be the result of government intervention in the form of zoning. The results of his study imply that countries with a higher share of residential land designated only for the construction of single family homes, *ceteris paribus*, have a higher homeownership rate. He suggests this to perhaps be the reason for countries with higher propensity for densification, like Switzerland and Germany, to have extremely low ownership rates, compared to their European counterparts. Other reasons have been proposed to explain Germany's comparatively low ownership rate (51% as of 2019,

Eurostat). These include, just to mention some: laws in favour of tenants (rent regulation), an extensive social housing sector, lack of subsidies and tax relief for homeowners and stable house prices (Voigtländer 2009). Therefore, Germany as a high-income modern state stands out to show that wealth does not always induce higher rates of owner-occupancy and besides demographics, legislative factors ought to be considered.

Analysing the effect of tax policies (Hilber omits mortgage market conditions from the current analysis), the non-taxation of imputed rents[3] appeared to have a strong positive effect on homeownership attainment. It must be noted that out of the 15 surveyed countries, only 4 (Belgium, Greece, Luxembourg and the Netherlands) still taxed imputed rents and the rest had abstained from including imputed rents in their tax base.

In conclusion, as evidenced by the discussion above, several factors are at play when analysing the dynamics of homeownership rates. Previous studies have explored a variety of factors associated with tenure choice and offered several, sometimes contradicting, explanations for cross-country and regional differences. Past findings, introduced in this section, will be considered in the following chapter when choosing appropriate variables and methodology.

[3] An estimate of the amount a homeowner would pay to rent their own dwelling that they are currently occupying themselves. Takes into consideration the rates charged for similar housing units on the market that are rentals and not owner-occupied.

2. DATA AND METHODOLOGY

The following chapter aims to give an overview of the underlying data and chosen variables used in the model and answer the research questions posed and presented in Introduction. The methodology and software program used for executing the empirical analysis, will also be discussed.

2.1. Variables

The chosen variables used in the regression are based on the factors explored in previous studies (summarized in Chapter 1.3.). *A priori* expectations are made about the effect of each variable and are presented in the following subchapters as hypotheses 1 to 4. The dependent variable is the homeownership rate while mortgage interest rates, house price index, price to rent ratio and annual median equivalised net income are included as independent variables. Analysis in subgroups will be performed using data on household type and income groups. Due to data limitations, it is not possible to use quarterly data as only annual observations are available for the most important datasets, retrieved from the EU Statistics of Income and Living Conditions (EU-SILC) database. Thus, annual observations from the year 2006 to 2019 are used.

The data is obtained as follows: homeownership rates (total and by relevant subgroups) and annual median equivalised net income data is retrieved from the EU-SILC database, mortgage interest rates from the statistical database of Bank of Estonia and price to rent ratio as well as housing price indices are obtained from OECD database on housing prices. The variables are further discussed in subchapters 2.1.1 to 2.1.4 and the complete set of values for the data used is presented in Appendix 1.

2.1.1 Homeownership rate

The subject of the regression will be the homeownership rate in Estonia from the year 2006 to 2019 extracted from the EU-SILC database, more precisely from the table “Distribution of

population by tenure status, type of household and income group" (ilc_lvho02). Homeownership rate is defined as the percentage of total households that are owner-occupied. The rest of the households qualify as tenants. Households qualified as owner-occupied include both households with mortgage or a loan and those without. The dynamics of homeownership rate is displayed on Figure 1.

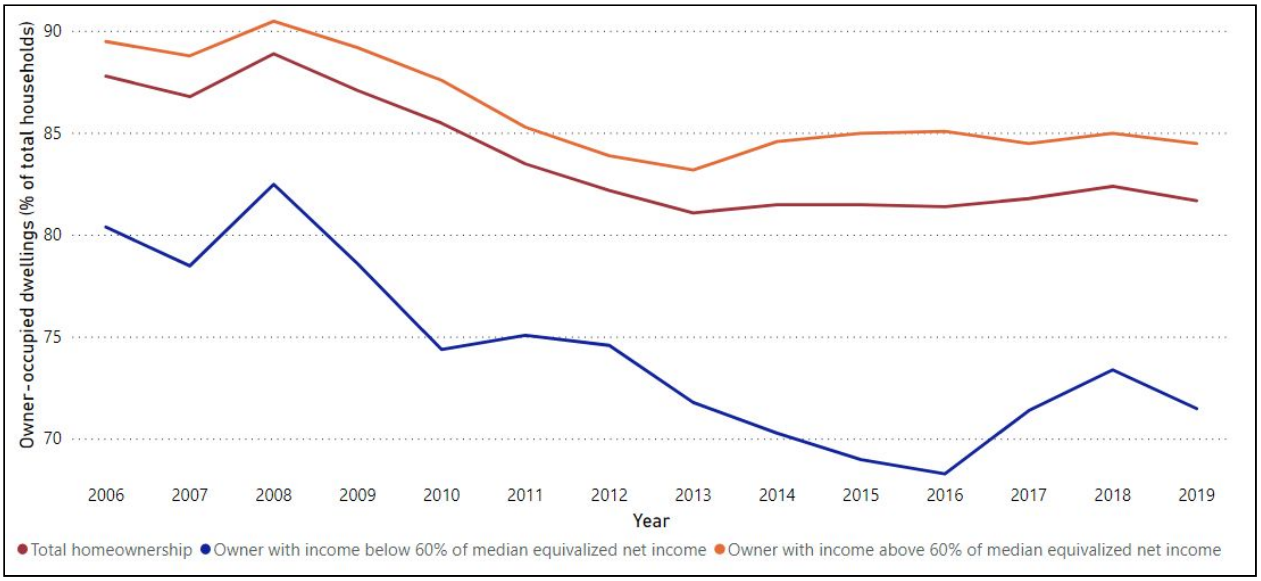

Figure 1. Homeownership rate in Estonia in total and by income groups.
Source: Eurostat, author's calculations

As can be seen from Figure 1, the proportion of households who are owner-occupiers started dropping during the worldwide economic crisis in 2008 and among households whose annual equivalised[4] income is above 60% of national median income, leveled off in 2013. Among lower income households, the decrease of homeownership rates was more dramatic and lasted until 2016. During the period studied, the drop in homeownership rate in the higher income cohort was around 7 percentage points while being exactly double that among households in the lower cohort (14 percentage points). The downturn in the economy led to a significant amount of households facing solvency problems and thus, a large share of the households who had used their primary residence as collateral, had their real estate assets foreclosed. In the period 2007-2013, over 4000 private real estate objects were foreclosed (Hankewitz 2013) and primary residences sold due to financial difficulties. These factors could account for some of the negative changes in homeownership rates.

[4] Defined by Eurostat as *"The equivalised disposable income is the total income of a household, after tax and other deductions, that is available for spending or saving, divided by the number of household members converted into equalised adults; household members are equalised or made equivalent by weighting each according to their age."*

When taking into account demographic characteristics, age is at first considered to be an appropriate factor to include in the model. However, due to data limitations it would only be possible to use aggregate mean/median age of the whole Estonian population or the population by age groups, which would be a rather crude attempt to capture the effect of age on homeownership. Therefore, the variable age was disregarded. Rental market regulations in Estonia are limited (almost nonexistent), so rental market conditions could not be assessed or included as a variable either.

To include the households' demographic characteristics, the type (size) of the household will alternatively be taken into account. EU-SILC provides a breakdown of total homeownership by household types of which the following will be separately modelled: total households, households consisting of a single adult, households with two adults and finally, those with two adults and two dependent children. This provides the opportunity to infer whether larger households are more prone to being homeowners, like deduced by for example Hilber (2007) and Garcia and Figuieira (2020) and whether the factors affecting varied households are different. Breakdown of ownership rates by selected[5] household types are presented in Table 1.

Table 1. Descriptive statistics of homeownership rates by household type (2006-2019)

	Proportion of total homeowners (%)	Minimum (%)	Maximum (%)
Single adult	11.14	10.3	12.9
Two adults	19.57	18.40	20.30
Two adults and two dependent children	13.20	12.40	13.90

Table 1 shows that the highest proportion of homeowners are those with two adults and no children (when considering other household types not included here, the largest share of homeowners were still households with two adults). The proportion of homeowners who are single adults is slightly smaller than households with two dependent children. This seems to contradict theory, that larger households are more likely to own their primary residence. Regression analysis by household type will be performed in chapter 3.3.2.

[5] Not all household types defined by EU-SILC are included.

2.1.2 Mortgage interest rates.

The impact of mortgage market conditions and access to credit have previously been studied for example by Chambers *et al.* (2009) and Andrews and Sánchez (2011). Average long-term Euro loan weighted annual interest rates of housing loans - from the statistics of Bank of Estonia - were included in the selection of variables. Housing loans are either mortgages or loans granted for rebuilding property. Logically, it is assumed that lower interest rates lead to higher lending activity and hence, higher rates of homeownership due to the availability of financing. Mortgage interest deductibility was already in place prior to the period studied and therefore it was not possible to assess its impact on borrowing activity or ownership rates. Figure 2 presents the average annual interest rates of housing loans for the time span 2006-2019.

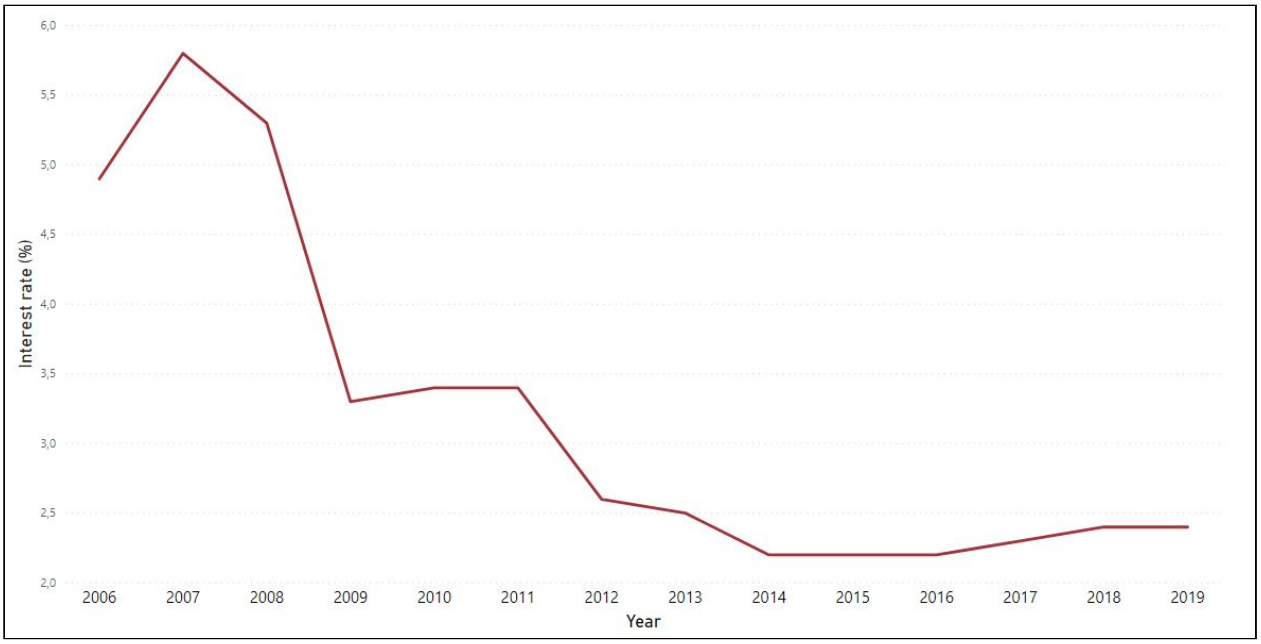

Figure 2. Average weighted annual interest rates of housing loans in Estonia from 2006 to 2019. Source: Bank of Estonia, author's calculations

During the period studied, mortgage interest rates reached their peak in 2007 and plummeted from 5.8 percent in 2007 to 3.2 percent by the end of 2009. The rates continued to drop but have been mostly stable from 2014, around 2.2 percent to 2.4 percent. The drop that escalated during the recession was caused by the change in EURIBOR, which plunged to zero by 2015 and has continued to have a negative value ever since[6]. This means that as long as the 6-month EURIBOR is zero or below zero, the only interest payable for a mortgage is the bank margin, which makes the cost of borrowing lower. Hypothesis 1 is proposed:

[6] A policy adopted by the European Central bank to boost lending to businesses and consumers.

H1: Mortgage interest rates have a negative relationship with homeownership rates

2.1.3 Income

Income could arguably be one of the most influential factors in determining tenure status and the propensity to own. Median equivalised net income (weighted anual income per household member) will be retrieved from Eurostat and included in the model as an independent variable. Considering the findings of Struyik and Marshall (1974) and Segal and Sullivan (1998) about the fact that increments in income affect lower earning households more in their tenure choice, different income groups will be analysed. The subgroups are households whose equivalised income exceeds 60% of the national median income and a lower cohort of those whose equivalised income is below that level. Table 2. shows average values for both groups.

Table 2. Homeownership by income groups

	Average homeownership rate (%)	Average annual equivalised income (EUR)
Income > 60% of national median	86.19	8 225.7
Income < 60% of national median	74.27	3 323.8

Source: Eurostat (EU-SILC), authors calculations

The total and the two subgroups of households are also presented in Figure 3. By the definition of Eurostat, *"anyone with an equivalised income of less than 60% of the national median is considered to be at risk of poverty"*. Thus, separate analysis will be performed to assess whether the homeownership rate among households in the lower end of the income spectrum have a higher responsiveness to changes in income.

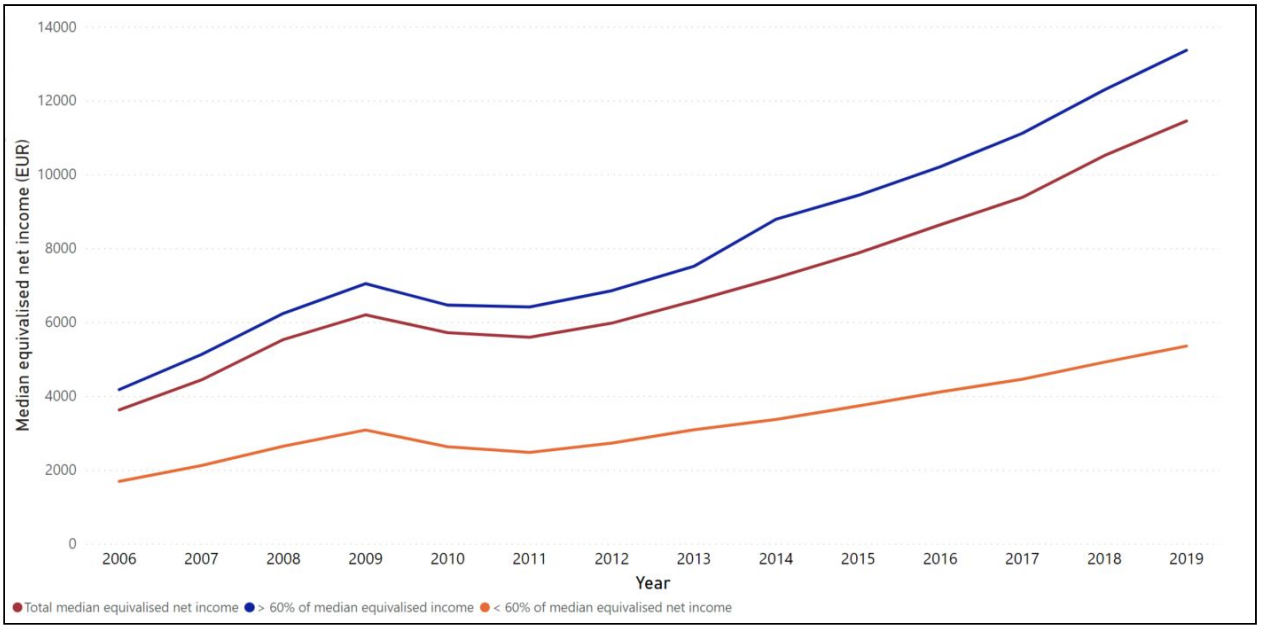

Figure 3. Annual median equivalised net income in Estonia from 2006 to 2019 by income groups.
Source: Eurostat, author's calculations.

As can be observed from Figure 3, there are significant differences between the upper and lower cohort in terms of earnings. A post-recession decrease in incomes is evident among all groups, yet from the slope of the graph we can see that earnings in the lower cohort increase at a slower pace than among high-earning households. Hypothesis 2 is proposed:

H2: Household income has a positive relationship with homeownership rates

2.1.4 House price index and price to rent ratio

To take into consideration the cost of owning a dwelling, two measures are included from the OECD housing prices database presented on Figures 4 and Figure 5. Firstly, the real house price index which by OECD's definition *" Covers the sales of both newly-built and existing dwellings. The real house price index is given by the ratio of the nominal house price index to the consumers' expenditure deflator in each country from the OECD national accounts database."* Secondly, price to rent ratio is included to measure the relative cost of owning. This ratio is retrieved by dividing the nominal house price index by the housing rent price index and can be used to measure the profitability of owning compared to renting. Higher (index value over a 100) price to rent ratio indicates owning a house is relatively more expensive than renting and vice versa. Both indices are seasonally adjusted and with base year 2015 (2015=100).

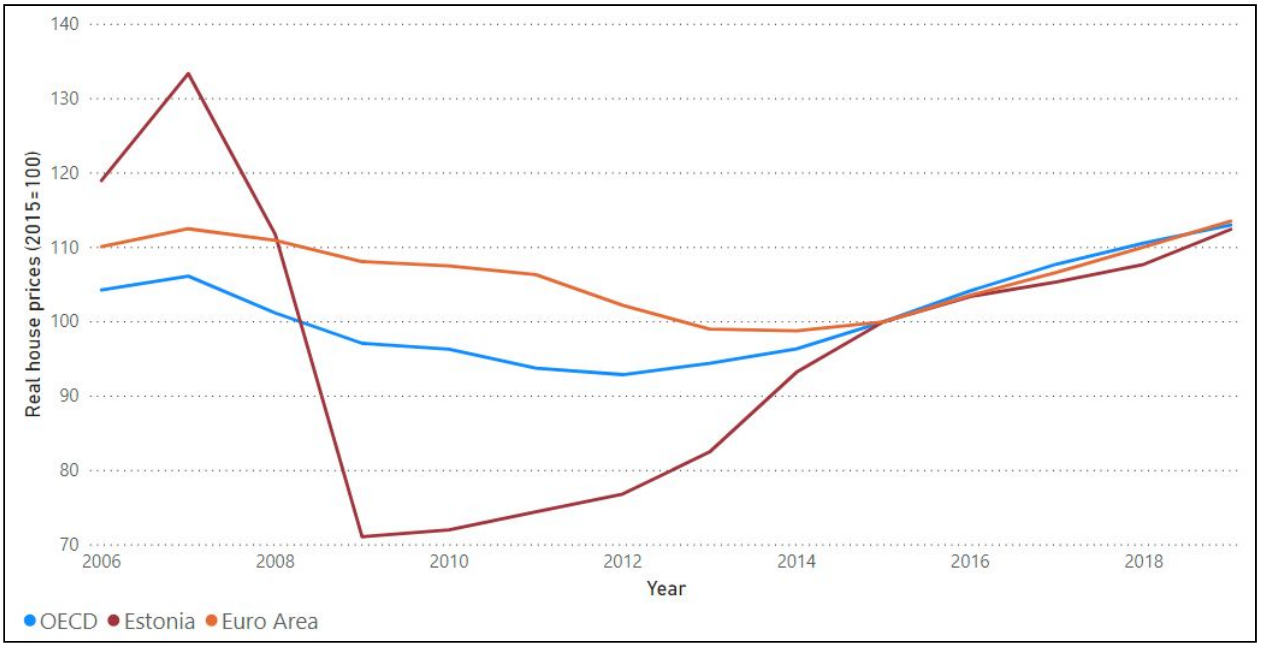

Figure 4. Real house price index in Estonia, Euro area and OECD countries from 2006 to 2019. (2015=100)
Source: OECD, author's calculations

By looking at Figures 4 and Figure 5, it emerges that Estonian housing prices - in the observed period - have been noticeably more volatile than in other OECD and Euro Area (EA) countries. Real house prices peaked in 2007 when the index reached a value of 133 and in the following two years plummeted to 71, nearly halving the value of housing stock. However, house prices have made a remarkable recovery since 2009 and by 2019, the index stood at 112. The drop in prices was a result of the 2007-2008 financial crisis, which was related to the inflated real estate prices and the so called real estate bubble bursting. Assumption about the effect of income on ownership rate is proposed in hypothesis 3:

H3: House prices have a negative relationship with homeownership rates

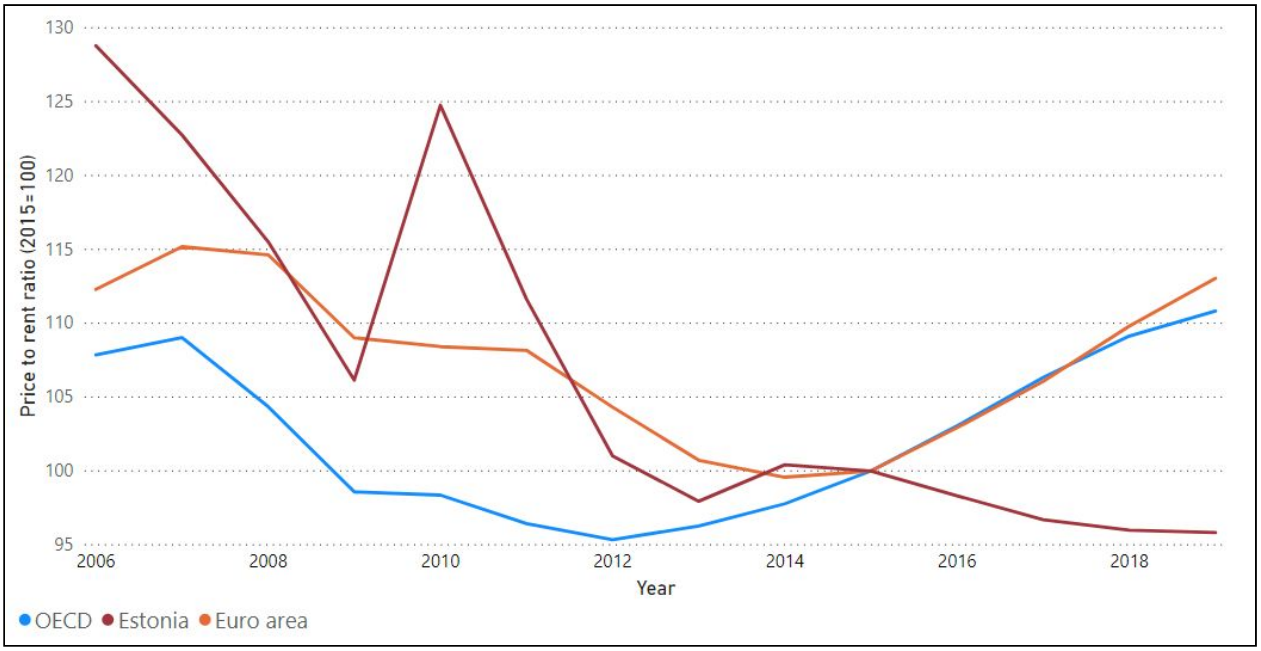

Figure 5. Price to rent ratio index in Estonia, Euro area and OECD countries from 2006 to 2019. (2015=100)
Source: OECD, author's calculations

The relative cost of owning, i.e. price to rent ratio has gone through dramatic changes during the period under study. Prior to the recession, choosing to rent instead of buying property was smarter due to the inflated house prices. Figure 5 shows that following the financial crisis, between the years 2012 and 2015 renting and owning offered more or less the same level of profitability (index value around 100). From 2015, owning has become relatively cheaper, because while both house and rent prices have constantly increased since 2009, the latter has done so at a much faster pace bringing the price to rent ratio below 100. Higher values of the price to rent ratio imply that owning a dwelling is relatively more expensive than renting. Hence, hypothesis 4 is proposed:

H4: Price to rent ratio has a negative relationship with homeownership rates

2.2 Methodology

2.2.1 Modelling a time series

The data used in the thesis is time series data with 14 annual observations. Multivariate regression analysis will be performed in econometric package Gretl (Gnu Regression, Econometrics and Time-series Library), using the ordinary least squares (OLS) method.

In order to model time series data using OLS, there are several assumptions that have to be fulfilled prior to modelling, the most important of which is that the time series has to be stationary. Stationary time series is a series that does not include a trend, for example interest rates tend to normally be stationary. Non-stationary time series are those that have a trend (e.g. real income, real GDP) and when used in regression analysis, can lead to spurious regression results (Sauga 2017, 581). To test whether the chosen variables are stationary, Augmented Dickey-Fuller (ADF) tests on each variable were performed in Gretl. The null hypothesis is that unit root is present and the time series is nonstationary.

The results of the ADF tests for the presence of unit root are as follows:

Stationary time series: interest rates, price to rent ratio, homeownership rates for total households, for those with two adults and those with two children;

Nonstationary time series: homeownership rates for households consisting of a single person, the time series' of homeownership rates by income subgroups and the time series for total income and house price index

The stationarity criterion for ownership rates among lower income households and for total income was achieved by taking the second-order difference of the time series. Stationarity for ownership rates for higher income households, the time series of house price index and homeownership rates for single households was achieved by taking the first difference. The results of the ADF tests for the variables are given in Appendices 2 to 13. In addition to Figures 1 to 5, an overview and descriptive statistics of the chosen variables as well as the ADF test results on a 0.05 significance level are given in Table 3.

Table 3. Descriptive statistics of the chosen variables.

	Variable	Mean	Minimum	Maximum	Standard deviation	ADF test p-value
1	Homeownership rate (total)	83.800	71.094	133.380	11.671	0.007
2	Average weighted housing loan interest rate	3.207	2.200	5.800	1.243	1.971e-13
3	First difference of house price index	-19.198	-40.663	4.692	22.774	0.015
4	Price to rent ratio (index)	106.845	95.823	128.797	11.671	3.796e-51
5	Second order difference of median equivalised net income	10.500	-1150.000	506.000	448.813	2.438e-05

Source: Eurostat and OECD, author's calculations

For variables 1 and 5, in addition to total values, analysis in subgroups will be performed in Chapter 3.3.

2.2.2 Ordinary Least Squares (OLS) method

To fulfill the aim of the thesis, regression analysis using the ordinary least squares (OLS) method is performed. After testing for stationarity and adjusting variables accordingly, linear regression model was deemed appropriate to model the selected data. Linear regression analysis using the OLS method returns a model that allows to estimate the relationship between a dependent variable and one or more independent variables. It does so by minimizing the sum of the squares in the difference between the observed and predicted values and generates a straight line as a result (Sauga 2017, 426). In regression analysis, it is also possible to interpret the magnitude and direction of impact that the independent variables have on the dependent variable. The model is easily adjustable and provides the opportunity to add or exclude variables to adjust the model if needed during execution. A linear regression line has an equation in the form (Sauga 2017):

$$y = b + a_1 x_1 + a_2 x_2 + ... + a_k x_k + \varepsilon \quad (1)$$

where

y is the dependent variable,

$x_1, x_2, ..., x_k$ are the independent variables,

$b, a_1, ..., a_k$ are parameters to be estimated;

ε is the error term

The equation for the initial model is presented in chapter 3.1.1 (Equation 2).

3. REGRESSION ANALYSIS AND RESULTS

3.1. The baseline model

The initial regression model was created in Gretl using the OLS method. The dependent variable is homeownership rate and the independent variables are:mortgage interest rates, second-order difference of annual median equivalised income, first difference of house price index and price to rent ratio (indexed values). Equation 2 presents the coefficients for the initial model:

$$own_{total} = 74.419 + 1.508INT - 0.001dd_{income} - 0.075\ d_{PRICE} + 0.042Price\_to\_rent \qquad (2)$$

The full report for the model is presented in Appendix 14. The significance level used throughout the analysis is 0.05. The model as a whole is statistically significant ($p = 3.63e-06$) and the coefficient of determination (R-squared) was 0.982, meaning that the model explained about 98 percent of the variance in homeownership rates. Excluding the constant, $p-value$ is the lowest for house price index ($p = 0.001$). However, both price to rent ratio and income variable turned out to be statistically insignificant. Hence, a correlation matrix was drawn up to remove insignificant variables, starting from the one least correlated to y. The correlation matrix can be found in Appendix 15. The variable with the weakest relationship to homeownership rate was the price to rent ratio ($r = 0.075$), which was removed and a new model was estimated (Model 2 in Appendix 16). The equation for the adjusted model is presented below, Equation 3:

$$own_{total} = 77.705 + 1.897\ INT - 0.001dd_{income} - 0.067d\_Price \qquad (3)$$

In the adjusted model, all variables and the model as a whole were statistically significant. R-squared decreased slightly from 0.981 in Model 1 to 0.978 in Model 2. The adjusted R-squared stayed almost the same (0.970 compared to 0.972 in Model 1). Therefore, by disregarding price to rent ratio from the variables, it is possible to get a statistically significant model with statistically significant variables, yet not lose any explanatory power.

3.2 Testing the adjusted regression model

In the following chapter, the assumptions of the classic linear regression model are tested for Model 2. Gujarati (2004) has proposed 10 assumptions[7] of which the following 5 were tested:

1) The mean value of residuals is zero;
2) Normal distribution of residuals;
3) Homoskedasticity;
4) No autocorrelation between residuals
5) No perfect multicollinearity

If one or more of the mentioned assumptions are not satisfied, any statistical inference about the dependent variable could be flawed and the interpretation of the regression estimates would not be valid.

The first assumption means that the impact of unknown independent variables, that are not in the model, on the dependent variable y is in total zero. This assumption is automatically satisfied by including a constant (Sauga 2017). Since Model 2 includes a constant, the first assumption will not be separately tested.

The second assumption is the normal distribution of residuals. Appendix 16 presents the full report for Model 2, including the results of Doornik-Hansen test for normality of residuals. Residuals are normally distributed ($p = 0.306$) and thus, the second assumption is satisfied.

The next criterion to be fulfilled is homoskedasticity, meaning that the residuals have a constant dispersion and are independent of exogenous variables. The complementary notion is called heteroskedasticity, in which case the standard errors in the model are wrong and cannot be relied upon. White's test in Gretl was used to test for possible heteroskedasticity. According to the test report available in Appendix 16, heteroskedasticity was not present ($p = 0.220 > 0.05$).

Autocorrelation is a phenomenon where the residuals in the regression are not independent and identically distributed (Sauga 2017) which leads to successive points in the time series being correlated to each other. Similarly to heteroskedasticity mentioned above, this can lead to wrong standard errors and a falsely estimated model. Since the sample size in the regression is

[7] The selection of assumptions to be tested is based on their importance (effect on the model). Some of the ones proposed in literature are already fulfilled, for example stationarity of the data or that the number of observations exceeds the number of parameters. Therefore, five out of ten will be discussed.

relatively small ($n = 14$), Breusch-Godfrey test (better suited for smaller samples) was chosen over Durbin-Watson's test to assess whether autocorrelation is present. The p-value of the test statistic LMF was 0.083 which exceeds 0.05. Therefore, the null hypothesis is accepted: no autocorrelation and assumption four is satisfied. Detailed results of the Breusch-Godfrey test are available in Appendix 16.

In order to see whether the last assumption is satisfied, each variable was tested for multicollinearity, which is present when two or more independent variables in the regression are correlated. Perfect multicollinearity exists when one independent variable has a linear relationship with another independent variable. It is difficult, especially in economics, to find variables that have no relation at all. Hence, when referring to multicollinearity in econometrics, approximate multicollinearity is usually meant (Sauga 2017, 486). When multicollinearity is present, the model may contrast theory and go against logic (*Ibid.*, 490). Variance Inflation Factor (VIF) was used to assess whether multicollinearity exists in the model. VIF values over 10 indicate possible multicollinearity. All VIF values stayed significantly under 10 and no multicollinearity was identified.

Finally Ramsey's RESET test was performed to confirm whether the shape of the model is correct. Null hypothesis was accepted ($p = 0.361$), the shape of the model is correct. Taking into account the above tests and satisfactory results, Model 2 can be used to make reliable predictions and estimates. Implications of Model 2 are discussed in Chapter 3.4.

3.3 Subgroup models

Presumably, the relationships estimated in Model 2 should hold for subgroup models as well. Previous literature has noted household size to play a role in the probability to own (e.g. Hilber 2007, Garcia, Figueira 2020) and hence, homeownership rates of different household types will be analysed in separate models. Homeownership rates by household types are presented in Table 1 in Chapter 2.1.1.

The impact of increments in income has been found to vary between different income cohorts. Struyik and Marshall (1974) and Segal and Sullivan (1998) noted that households in the lower end of the income spectrum have a higher responsiveness to changes in earnings and that homeownership rates among wealthier households are less sensitive to income. Therefore, homeownership rates for households whose equivalised net income is below 60% of the national

median income and ownership rates for households whose income exceeds that threshold will be modelled separately. Refer to Table 2 in Chapter 2.1.3 for a breakdown of homeownership rates by income groups.

The method will be the same (OLS) as with Model 2. The same tests for assumptions will be performed as well, but at this point discussed in less detail. An overview of the assumptions tested is given in Chapter 3.2.

3.3.1 Income groups model

The first model to be analysed is for lower income households (Model 3). The dependent variable is the second-order difference of homeownership rate for households whose annual equivalised income is below 60% of the national median net income. Independent variables are second-order difference of income (median equivalised income of households earning less than 60% of the national median), first difference of house price index, mortgage interest rates and price to rent ratio. The first estimated model is not statistically significant ($p = 0.106$) and the only statistically significant variable was house price index. Insignificant variables are removed according to the correlation matrix presented in Appendix 17. Both income and price to rent ratio must be excluded from the model due to being insignificant. Equation 4 presents Model 3:

$$dd\_own_{low} = -10.3504 + 3.771\, INT + 0.228\, d\_PRICE \qquad (4)$$

The R-squared for Model 3 is 0.586 and the p-value for overall significance is 0.019. The model is statistically significant but explains only about half (59%) of the variance in ownership rates among lower income households. The signs of the coefficients are opposite of what is expected, as increases in prices and interest rates are expected to have a negative relationship to homeownership rates. Residuals are normally distributed, heteroskedasticity, autocorrelation or multicollinearity are not identified (see Appendix 18 for the full report). Ramsey's RESET test, however, indicates that the model might miss important variable(s). Implications of model 3 are discussed in Chapter 3.4

The same method as with lower income households is used to estimate the model for homeownership rates in the higher income cohort. The dependent variable is the homeownership rate for households with income above 60% of median equivalised net income. The independent variables are the second-order difference of income (median equivalised income of households earning more than 60% of the national median), first

difference of house price index, mortgage interest rates and price to rent ratio. Surprisingly, the estimated model as a whole as well as all the variables are statistically insignificant. Examining the correlation matrix between the variables (Appendix 19), it is evident that none of the variables have a strong correlation with the dependent variable. Price to rent ratio has the strongest correlation to the dependent variable ($r = -0.242$) which is still relatively low. The low correlation coefficients in the correlation matrix already indicate that a model with relationships that are weak will possibly turn out to be insignificant. Combinations of regressors were tried, but the model and the variables remained insignificant nevertheless. The limitations of the model are discussed in Chapter 3.4.

3.3.2 Household types model

Three separate household types are attempted to be analysed:

1) Households consisting of a single person (nonstationary, first difference taken)
2) Households consisting of two adults (stationary)
3) Households with two dependent children (stationary)

Ownership rates of different household types are set as dependent variables while the independent variables are the same as in the initial adjusted model (Model 2). OLS models for all of the three subgroups present the same issue as the model in chapter 3.3.1 for higher income households. None of the variables are significant and neither the model as a whole. Again, all of the correlation matrices show very weak relationships, with r values staying mostly under 0.3. Different combinations of variables are tried in the OLS model, yet none of them were significant, neither on their own nor in a combination with others. The model as a whole stayed statistically insignificant.

3.4. Analysis of results

The aim of the regression analysis performed was to determine which chosen independent variables have an impact on homeownership rates and whether this impact is positive or negative. A regression model was estimated using the OLS method for total homeownership rate in Estonia in the period from 2006 to 2019. The dependent variable was homeownership rates

and the independent variables were average weighted mortgage interest rates, second-order difference of annual median equivalised net income, first difference of house price index and price to rent ratio (indexed values).

In the initial model, price to rent ratio was not statistically significant and was excluded. Therefore, hypothesis 4 which stated that price to rent ratio has a negative relationship with homeownership rates, was rejected. Price to rent ratio had no significant effect on ownership rates. The insignificance of price to rent ratio in the estimated model could be explained by the fact that people, in economic decision making, tend to focus more on absolute monetary values than relative (sometimes abstract) prices and ratios. Therefore, housing prices in general should affect tenure choice decisions more than the relative price of owning versus renting.

After the price to rent ratio was excluded from the model, all the variables and the model as a whole were statistically significant with a high explanatory power $R-squared = 0.978$. Equation 3 presents the adjusted Model 2:

$$own_{total} = 77.705 + 1.897\ INT - 0.001 dd_{income} - 0.067 d\_Price \qquad (3)$$

As Model 2 was successfully tested against the assumptions of classical linear regression, implications can be made. Firstly, increases in house prices have a negative relationship with homeownership. Therefore, hypothesis 3 is accepted. The model indicates that as the house price index increases by one unit, total homeownership rates decrease by 0.07 percentage points. The small coefficient is logical, as homeownership is observed as rather inelastic and does not respond rapidly to changes in demand.

Secondly, average weighted mortgage interest rates were estimated to have a positive relationship with homeownership rates. Hence, hypothesis 1, in which a negative relationship was expected, is rejected. The model suggests that as mortgage interest rates increase by one percentage point, homeownership rates increase by 1.90 percentage points. While this finding contradicts logic, a possible explanation can be found when analysing the dynamics of both total homeownership and interest rates. The time series used in the regression includes a period of economic downturn, from 2007 to 2009/2010. During a recession, the drop and a slow recovery of economic indicators is expected. In that period, both interest and homeownership rates dropped, but likely due to different reasons. While interest rates were lowered and kept at extremely low levels as a policy of the ECB, homeownership rates decreased due to sales and

foreclosures, brought upon by households' solvency problems and financial difficulties. As the length of the time series is short and sample size relatively small, long term inference about interest rates can not be drawn. Especially since there were significant fluctuations in most economic indicators during the recession and in the wake of it. A longer time series or quarterly data may potentially improve the accuracy of the model.

Income was found to have a negative relationship with homeownership rates. Thus, hypothesis 2 is rejected. The coefficient for income (0.001) can be interpreted as follows: as a household's annual equivalised income increases by 1 euro, ownership rates decrease by 0.001 percentage points. In other words, a 1 000 euro increase in equivalised annual income decreases homeownership rates by 1 percentage point. The reason for the negative coefficient, that conflicts with theory, could be that the relationship between income and homeownership is in fact not linear. The nonlinearity of the relationship between income and homeownership rates was proposed already by Struyk and Marshall (1974) and Segal and Sullivan (1998). The former authors used the quadratic form of permanent income in their model and confirmed the relationship between income and ownership rates is non linear. This implies that the model could be developed further and adjusted for a non linear regression, which might offer results that are more realistic and reliable. Alternative measures of income or wealth (e.g. GDP per capita) could be explored as well.

Analysis in subgroups resulted mostly in models that were not statistically significant and could not be used for inference. The only statistically significant model that offered some explanatory power was Model 3, homeownership rates of low-income households was modelled. The equation of Model 3 is described below:

$$dd\_own_{low} = -10.3504 + 3.771\ INT + 0.228\ d\_PRICE \qquad (4)$$

The coefficient signs, as opposed to theory, indicate that homeownership rates among low income households have a positive relationship with prices and interest rates. The low explanatory power $(R-squared = 0.586)$ indicates that an important variable might me missing and more information is needed.

The model for ownership rates among relatively higher income households was not statistically significant and neither were none of the variables. Potential reasons behind these results might be explained by fact that medium and high income households have accumulated enough wealth and hence, their tenure choice decision is affected less by increases in income, interest rates and

prices. Purely speculative explanations can be suggested for the insignificant relationship between the chosen variables. No reliable inference can be made, as the attempt to estimate a model from the chosen variables using OLS failed for this sample.

Lastly, the homeownership rates of three types of households (different number of members) were analysed in separate models as a robustness check. None of the models were significant and had no explanatory power. The correlation between the dependent and independent variables was relatively weak, which implies that the factors affecting tenure choice decisions of different kinds of households vary. The results of the model for total households did not hold in the subgroup models. This indicates that more information is needed. Important explanatory variables might be missing from the regression and subgroup analysis may require a different approach than the analysis for total households.

CONCLUSION

Homeownership rates have been subject to extensive research in the recent decades. As discussed in the literature review of this thesis, homeownership has been linked to a host of socio-economic benefits both on micro level and for the society as a whole. The purpose of the thesis was to explore the determinants of homeownership rate in Estonia and to evaluate if and to what extent homeownership rate is influenced by mortgage market conditions, selected socio-economic and demographic factors.

A research question "Is there a statistically significant relationship between homeownership and one or more of the studied independent variables (income, mortgage interest rates, house price index, price to rent ratio)?" was proposed. In addition, four hypotheses were established:

H1: Mortgage interest rates have a negative relationship with homeownership rates
H2: Household income has a positive relationship with homeownership rates
H3: House prices have a negative relationship with homeownership rates
H4: Price to rent ratio has a negative relationship with homeownership rates

In order to answer the research question and test the hypotheses, regression analysis was performed using the OLS method in Gretl software. Time series data with annual observations from 2006 to 2019 was modelled. The dependent variable was homeownership rate and the independent variables were average weighted mortgage interest rates, house price index, price to rent ratio and annual median equivalised net income. Analysis in subgroups was performed using data on household type and income groups. Before modelling, the time series were tested for stationarity and adjusted accordingly. Assumptions of a classic linear regression model were tested and fulfilled.

The results of the regression for the baseline model showed that the mortgage interest rates, income and house prices were statistically significant. Price to rent ratio was statistically insignificant and was thus removed from the model. The adjusted model was statistically significant and had a high explanatory power of 98%. Assumptions of a classic linear regression

model were tested and satisfactory results achieved. Therefore, the model could be used to make reliable predictions and estimates.

However, out of the four hypotheses proposed, only hypothesis no 3 was accepted - house prices have a negative relationship with homeownership rates. This finding confirms what has been noted in previous studies that, as house prices increase, homeownership rates decrease

Mortgage interest rates and income had a negative relationship, which contradicted with previous studies analysing the effect of developments in the mortgage market conditions. The positive relationship between ownership rates and mortgage interest rates could possibly be explained by the fact that during the economic recession (2007-2009) and in the wake of it, both indicators dropped, but due to different reasons. Interest rates were lowered by the ECB as a policy to boost lending, yet homeownership rates dropped due to households facing solvency problems, foreclosures and financial difficulties. Therefore, while the two variables were positively related, this does not indicate a causal relationship between them.

The negative effect of income on ownership rates could be due to the fact that the relationship may in fact be nonlinear. The negative sign could indicate that the shape of the model is wrong. The regression model could be developed further and adjusted for a non linear regression (for example, use the square of income), which might offer results that are more realistic and reliable. Alternative measures of income or wealth (e.g. GDP per capita, the square of income) could be explored as well.

Models using subsamples of homeownership rates by income groups and household type were estimated and used for robustness check. Analysis in subgroups resulted mostly in models that were not statistically significant and could not be used for inference. The only statistically significant model that offered some explanatory power was the model where homeownership rates for low-income households were analysed. However, the explanatory power of the model was relatively low (59%), meaning that important variables were most likely not included in the model. Therefore, for the most part the models for subsamples could not be used for inference and no reliable conclusions could be drawn.

The limitations of the analysis were definitely the small sample size (n=14) and annual observations. Sample size can be increased by using quarterly data or a different time period. Alternative models and other variables could be explored, as nonlinear models may result in different outcomes and offer results that do not contradict previous literature on tenure choice..

The fact that models for ownership rates by income groups and household type were nonsignificant offers the opportunity for further research. Different demographic groups could be explored separately as the determinants of homeownership across households are most likely not uniform.

REFERENCES

Andrews, D., & Sánchez. A. (2011). The Evolution of Homeownership Rates in Selected OECD Countries: Demographic and Public Policy Influences. *OECD Journal: Economic Studies, 2011*(1), 197-233. https://doi.org/10.1787/eco_studies-v2011-1-en

Bourassa, S. C., Haurin, D. R., Hendershott, P. H., & Hoesly, M. (2011). Determinants of the Homeownership Rate: A Survey of Recent Contributions. *Swiss Finance Institute Research Paper, 11-49.* http://dx.doi.org/10.2139/ssrn.1961189

Capozza, D. R., Green, R. K., & Hendershott P. H. (1997). Taxes, Mortgage Borrowing, and Residential Land Prices. *Economic Effects of Fundamental Tax Reform*

Carliner, G. (1974). Determinants of Home Ownership. *Land Economics, 50*(2), 109-119. doi:10.2307/3145360

Chambers, M., Garriga, C., & Schlagenhauf D. E. (2009). Accounting for Changes in the Homeownership Rate. *International Economic Review, 50*(3), 677-726. https://doi.org/10.1111/j.1468-2354.2009.00544.x

Coulson, N. E. (2002). Housing Policy and the Social Benefits of Home Ownership. Federal Reserve Bank of Philadelphia, Business Review, Q2, 7-16. https://www.phil.frb.org/-/media/research-and-Data/publications/business-review/2002/q2/brq202ec.pdf

Coulson, N. E. (2002). Why Are Hispanic- and Asian-American Homeownership Rates So Low?: Immigration and Other Factors. *Journal of Urban Economics*, 45(2), 209-227. https://doi.org/10.1006/juec.1998.2094

Dietz, R. D., & Haurin, D. R. (2003). The Social and Private Micro-Level Consequences of Homeownership. Journal of Urban Economics, 54(3), 401-450. https://doi.org/10.1016/S0094-1190(03)00080-9

DiPasquale, D., & Glaeser, E. L. (1999). Incentives and Social Capital: Are Homeowners Better Citizens? Journal of Urban Economics, 45(2), 354-384. https://doi.org/10.1006/juec.1998.2098

Eesti Pank. (2020). Average long-term Euro loan weighted interest rates granted to non-fianncial corporations and housing loans, and 6-month EURIBOR [Data file]. Retrieved from: https://statistika.eestipank.ee/#/en/p/979/r/1654/1503

Eesti Statistika. (2011). Conventional dwelling by time of construction of building, owner and location [Data file]. Retrieved from: http://andmebaas.stat.ee/Index.aspx?lang=en&SubSessionId=78e9d7af-6418-4cec-a303-91bc77360200&themetreeid=-200

Eurostat. (2020). Distribution of population by tenure status, type of household and income group [Data file]. Retrieved from: https://ec.europa.eu/eurostat/databrowser/view/ilc_lvho02/default/table?lang=en

Eurostat. (2020). Mean and median income by household type [Data file]. Retrieved from: https://ec.europa.eu/eurostat/databrowser/view/ilc_di04/default/table?lang=en

Evans, G. W., Wells, N. M., & Moch, A. (2003). Housing and Mental Health: A Review of the Evidence and a Methodological and Conceptual Critique. Journal of Social Issues, 59(3), 475-500. https://doi.org/10.1111/1540-4560.00074

Fisher, L. M., & Jaffe A. J. (2003). Determinants of International Home Ownership Rates. *Housing Finance International, 18*(1)*,* 34-42. https://www.researchgate.net/publication/245900864_Determinants_of_International_Home_Ownership_Rates

Garcia, M. T. M., & Figueira, R. (2020). Determinants of Homeownership in Europe - an Empirical Analysis based on SHARE. *International Journal of Housing Markets and Analysis.* Advance online publication. doi:10.1108/IJHMA-12-2019-0120

Gatzlaff, D. H., Green, R. K., & Ling, D. C. (1998). Cross-Tenure Differences in Home Maintenance and Appreciation. Land Economics, 74(3), 328-342. https://doi.org/10.2307/3147116

Glaeser, E. L., & Shapiro, J. M. (2003). The Benefits of the Home Mortgage Interest Deduction. *Tax Policy and the Economy, 17*, 37-82. https://www.jstor.org/stable/20140504?refreqid=excelsior%3Af4229f9db57b262de7ce528ec55cfe3e&seq=1

Green, R. K., &White, M. J. (1997). Measuring the Benefits of Homeowning: Effects on Children. *Journal of Urban Economics*, 41(3), 441-461. https://doi.org/10.1006/juec.1996.2010

Gujarati, D. N. (2004). *Basic Econometrics.* (4th Edition). Tata McGraw Hill.

Gwin, C. R., & Ong, S. (2008). Do we really understand homeownership rates? An international study. *International Journal of Housing Markets and Analysis, 1*(1), 52-67. doi:10.1108/17538270810861157

Gyourko, J. E., & Linneman, P. D. (1997). The Changing Influences of Education, Income, Family Structure, and Race on Homeownership by Age over Time. *Journal of Housing Research 8*(1)*,* 1-25. https://www.jstor.org/stable/24833630?seq=1#metadata_info_tab_contents

Gyourko, J. E., Linneman, P. D., & Wachter, S. M. (1999). Analysing the Relationships Among Race, Wealth, and Home Ownership in America. *Journal of Housing Economics 8*(2)*,* 63-89. doi:10.1006/JHEC.1999.0242

Hankewitz, G. (2013, February 7). *Foreclosures peak last year*. Postimees. https://news.postimees.ee/1129916/foreclosures-peak-last-year

Haurin, D. R., Parcel, T. L., & Haurin R. J. (2001). The Impact of Homeownership on Child Outcomes. *Joint Center for Housing Studies of Harvard University.* https://www.jchs.harvard.edu/research-areas/working-papers/impact-homeownership-child-outcomes

Hilber, C. A. L. (2007). The Determinants of Homeownership across Europe: Panel Data Evidence. http://personal.lse.ac.uk/hilber/hilber_wp/hilber_2007_ho_europe_narsc.pdf

Kährik, A., Lux, M., Kõre, J., Hendrikson, M., & Allsaar, I. (2004). *Eluasemepoliitika üleminekuriikides.* PRAXIS Poliitikauuringute keskus. http://www.digar.ee/id/nlib-digar:713

Lauridsen, J., Nannerup, N., & Skak, M. (2006). *Explaining home ownership rates in Danish municipalities.* In J. Doling, & M. Elsinga (Eds.), *Home Ownership: Getting in, getting from, getting out. Part II* (59-17). IOS Press. https://www.researchgate.net/publication/242143674_Determinants_of_Homeownership_in_Denmark

Lauridsen, J., Nannerup, N., & Skak, M. (2013). Does Owner-Occupied Housing Affect Neighbourhood Crime? *SSRN, Electronic Journal.* DOI: 10.2139/ssrn.2362885

Marcuse, P. (1996). Privatization and its Discontents: Property Rights in Land and Housing in the Transition in Eastern Europe. In G. Andrusz, M. Harloe, & I. Szelenyi (Eds.), *Cities After Socialism: Urban and Regional Change and Conflict in Post-Socialist Societies* (pp. 119-191). Wiley-Blackwell. DOI: 10.1002/9780470712733.ch5

OECD. (2020). Housing Prices [Data file]. Retrieved from: https://data.oecd.org/price/housing-prices.htm

Oja, B. (2020, May 23). *Kinnisvarahinnad võivad sügisest langeda.* Eesti Rahvusringhääling. https://www.err.ee/1093644/kinnisvarahinnad-voivad-sugisest-langeda

Purju, A. (1996). The Political Economy of Privatisation in Estonia. CERT Discussion Papers, 9602. https://www.researchgate.net/publication/5101654_The_Political_Economy_of_Privatisation_In_Estonia

Rohe, W. M., & Lindblad, M. (2013). Reexamining the Social Benefits of Homeownership after the Housing Crisis. *Joint Center for Housing Studies of Harvard University.* https://www.jchs.harvard.edu/sites/default/files/hbtl-04.pdf

Rohe, W. M., & Stewart, L. S. (2010). Homeownership and Neighborhood Stability. *Housing Policy Debate, 7*(1), 37-81. https://doi.org/10.1080/10511482.1996.9521213

Rosenblad, Y. (2011). Tööturg majanduskriisi teises pooles. *Quarterly Bulletin of Statistics Estonia, 11*(2), 65-90. https://www.stat.ee/sites/default/files/2020-07/kvartalikiri_nr_2-11_veebi.pdf

Sauga, A. (2017). *Statistikaõpik majanduseriala üliõpilastele.* TTÜ Kirjastus. https://digikogu.taltech.ee/et/Item/76e80221-4d41-40e6-b395-a6ef16609482

Segal, L. M., & Sullivan, D. G. (1998). Trends in Homeownership: Race, Demographics, and Income. Economic Perspectives, 22(2), 53-72. https://www.researchgate.net/profile/Daniel_Sullivan9/publication/23529690_Trends_in_Homeownership_Race_Demographics_and_Income/links/55eaf7e708ae21d099c5dd2e.pdf

Struyk, R. J., & Marshall, S. (1974). The Determinants of Household Home Ownership. *Urban Studies 11*(3)*,* 289-299. https://www.jstor.org/stable/43081312?seq=1

Tuvikene, T. (2019). Between Community and Private Ownership in Centrally Planned Residential Space: Governing Parking in Socialist Housing Estates. In D. B. Hess, & T.

Tuvikene (Eds.), Housing Estates in the Baltic Countries. The Legacy of Central Planning in Estonia, Latvia and Lithuania (321-336). Springer International Publishing. DOI: 10.1007/978-3-030-23392-1_15

Valkiainen, V. (2005, January 27). *Erastamisväärtpabereid saab kasutada kuni 2008. Aastani.* Rahandusministeerium. https://www.rahandusministeerium.ee/et/uudised/erastamisvaartpabereid-saab-kasutada-kuni-2008-aastani

Voigtländer, M. (2009). Why is the German Homeownership Rate so Low? Housing Studies, 24(3), 355-372. https://doi.org/10.1080/02673030902875011

Yun, L., & Evangelou, N. (2016). Social Benefits of Homeownership and Stable Housing. *National Association of Realtors.* https://cdn.nar.realtor/sites/default/files/documents/Homeownership-Stable-Housing2016.pdf

APPENDICES

Appendix 1. Values of variables used in the regression (2006-2019)

Date	Mortgage interest rates (%)	Real house price index	Price to rent ratio	HO rate (total HH)	HO rate (HH single)	HO rate (HH 2 adults)	HO rate (HH 2 adults 2 dep. children)	HO rate (Income > 60% median equivalised net income)	HO rate (Income < 60% median equivalised net income)	Above 60% of median equivalised net income	Below 60% of median equivalised net income	Total median equivalised net income
31/12/2006	4,900	118,990	128,797	87,800	10,600	18,400	12,600	89,500	80,400	4185,000	1702,000	3636,000
31/12/2007	5,800	133,380	122,746	86,800	10,300	19,300	12,900	88,800	78,500	5134,000	2130,000	4447,000
31/12/2008	5,300	111,757	115,520	88,900	11,500	20,100	13,800	90,500	82,500	6249,000	2652,000	5541,000
31/12/2009	3,300	71,094	106,154	87,100	10,800	20,200	13,700	89,200	78,600	7054,000	3091,000	6209,000
31/12/2010	3,400	72,032	124,762	85,500	10,800	19,700	13,200	87,600	74,400	6472,000	2638,000	5727,000
31/12/2011	3,400	74,431	111,637	83,500	10,700	19,600	13,300	85,300	75,100	6422,000	2486,000	5603,000
31/12/2012	2,600	76,836	101,013	82,200	11,000	19,300	12,700	83,900	74,600	6864,000	2736,000	5985,000
31/12/2013	2,500	82,494	97,947	81,100	10,300	19,700	12,400	83,200	71,800	7524,000	3098,000	6583,000
31/12/2014	2,200	93,211	100,416	81,500	10,500	20,300	12,400	84,600	70,300	8798,000	3378,000	7210,000
31/12/2015	2,200	100,000	100,000	81,500	10,900	19,400	13,900	85,000	69,000	9443,000	3740,000	7882,000
31/12/2016	2,200	103,404	98,319	81,400	11,200	19,000	13,600	85,100	68,300	10217,000	4123,000	8647,000
31/12/2017	2,300	105,384	96,699	81,800	12,200	19,600	13,600	84,500	71,400	11124,000	4467,000	9389,000
31/12/2018	2,400	107,733	95,993	82,400	12,300	19,700	13,700	85,000	73,400	12302,000	4928,000	10521,000
31/12/2019	2,400	112,424	95,823	81,700	12,900	19,700	13,000	84,500	71,500	13372,000	5364,000	11458,000

Appendix 2. Testing for the presence of unit root in the time series for second order difference of median equivalised net income

```
k =  3: AIC = 102.991
k =  2: AIC = 107.581
k =  1: AIC = 111.383
k =  0: AIC = 112.510

  test with constant
  including 3 lags of (1-L)dd_Total_median_eqv_income
  model: (1-L)y = b0 + (a-1)*y(-1) + ... + e
  estimated value of (a - 1): -3.38694
  test statistic: tau_c(1) = -4.96225
  asymptotic p-value 2.438e-05
  1st-order autocorrelation coeff. for e: -0.214
  lagged differences: F(3, 3) = 5.958 [0.0884]

Augmented Dickey-Fuller regression
OLS, using observations 2012-2019 (T = 8)
Dependent variable: d_dd_Total_median_eqv_income

                  coefficient  std. error  t-ratio  p-value
  -------------------------------------------------------------
  const             376.018      95.5685     3.935   0.0292   **
  dd_Total_media~_1  −3.38694     0.682542  −4.962   2.44e-05 ***
  d_dd_Total_med~_1   1.08842     0.292421   3.722   0.0338   **
  d_dd_Total_med~_2   0.689763    0.211992   3.254   0.0474   **
  d_dd_Total_med~_3   0.256596    0.130987   1.959   0.1450

  AIC: 102.991   BIC: 103.388   HQC: 100.312
```

Appendix 3. Testing for the presence of unit root in the time series for income among households above 60% of median equivalised net income.

k = 3: AIC = 121.054
k = 2: AIC = 122.900
k = 1: AIC = 121.615
k = 0: AIC = 119.910

Augmented Dickey-Fuller test for dd_Above_60_median_income
testing down from 3 lags, criterion AIC
sample size 11
unit-root null hypothesis: a = 1

test without constant
including 0 lags of (1-L)dd_Above_60_median_income
model: (1-L)y = (a-1)*y(-1) + e
estimated value of (a - 1): -1.08569
test statistic: tau_nc(1) = -3.45382
p-value 0.002692
1st-order autocorrelation coeff. for e: -0.006

Dickey-Fuller regression
OLS, using observations 2009-2019 (T = 11)
Dependent variable: d_dd_Above_60_median_income

	coefficient	std. error	t-ratio	p-value	
dd_Above_60_me~_1	−1.08569	0.314343	−3.454	0.0027	***

AIC: 172.461 BIC: 172.859 HQC: 172.21

Appendix 4. Testing for the presence of unit root in the time series for income among households below 60% of median equivalised net income.

k = 3: AIC = 100.628
k = 2: AIC = 98.7720
k = 1: AIC = 96.8503
k = 0: AIC = 102.498

Augmented Dickey-Fuller test for dd_Below_60_median_income
testing down from 3 lags, criterion AIC
sample size 10
unit-root null hypothesis: a = 1

test without constant
including one lag of (1-L)dd_Below_60_median_income
model: (1-L)y = (a-1)*y(-1) + ... + e
estimated value of (a - 1): -1.49198
test statistic: tau_nc(1) = -3.21901
asymptotic p-value 0.001258
1st-order autocorrelation coeff. for e: -0.138

Augmented Dickey-Fuller regression
OLS, using observations 2010-2019 (T = 10)
Dependent variable: d_dd_Below_60_median_income

	coefficient	std. error	t-ratio	p-value	
dd_Below_60_me~_1	−1.49198	0.463490	−3.219	0.0013	***
d_dd_Below_60_~_1	0.423954	0.320040	1.325	0.2219	

AIC: 146.366 BIC: 146.971 HQC: 145.702

Appendix 5. Testing for the presence of unit root in the time series for average weighted interest rates for housing loans

k = 4: AIC = 2.42132
k = 3: AIC = 0.708222
k = 2: AIC = 0.230658
k = 1: AIC = 3.08206
k = 0: AIC = 1.26240

test with constant
including 2 lags of (1-L)INT
model: (1-L)y = b0 + (a-1)*y(-1) + ... + e
estimated value of (a - 1): -0.584452
test statistic: tau_c(1) = -8.13729
asymptotic p-value 1.971e-13
1st-order autocorrelation coeff. for e: 0.489
lagged differences: F(2, 7) = 8.480 [0.0135]

Augmented Dickey-Fuller regression
OLS, using observations 2009-2019 (T = 11)
Dependent variable: d_INT

	coefficient	std. error	t-ratio	p-value	
const	1.27619	0.215960	5.909	0.0006	***
INT_1	−0.584452	0.0718239	−8.137	1.97e-13	***
d_INT_1	−0.255013	0.104117	−2.449	0.0442	**
d_INT_2	−0.300186	0.0898557	−3.341	0.0124	**

AIC: -1.2056 BIC: 0.385976 HQC: -2.20887

Appendix 6. Testing for the presence of unit root in the time series for the first difference of house price index

k = 4: AIC = 42.9886
k = 3: AIC = 45.4478
k = 2: AIC = 44.9556
k = 1: AIC = 42.9935
k = 0: AIC = 41.7550

Augmented Dickey-Fuller test for d_PRICE
testing down from 4 lags, criterion AIC
sample size 12
unit-root null hypothesis: a = 1

test without constant
including 0 lags of (1-L)d_PRICE
model: (1-L)y = (a-1)*y(-1) + e
estimated value of (a - 1): -0.71404
test statistic: tau_nc(1) = -2.57508
p-value 0.01497
1st-order autocorrelation coeff. for e: 0.274

Dickey-Fuller regression
OLS, using observations 2008-2019 (T = 12)
Dependent variable: d_d_PRICE

	coefficient	std. error	t-ratio	p-value	
d_PRICE_1	−0.714040	0.277289	−2.575	0.0150	**

AIC: 98.373 BIC: 98.8579 HQC: 98.1934

Appendix 7. Testing for the presence of unit root in the time series for price to rent ratio (indexed values)

k = 4: AIC = 14.7245
k = 3: AIC = 15.6070
k = 2: AIC = 34.2295
k = 1: AIC = 35.4057
k = 0: AIC = 38.5100

test with constant
including 4 lags of (1-L)Price_to_rent
model: (1-L)y = b0 + (a-1)*y(-1) + ... + e
estimated value of (a - 1): -0.538219
test statistic: tau_c(1) = -22.7873
asymptotic p-value 3.796e-51
1st-order autocorrelation coeff. for e: -0.003
lagged differences: F(4, 3) = 24.887 [0.0123]

Augmented Dickey-Fuller regression
OLS, using observations 2011-2019 (T = 9)
Dependent variable: d_Price_to_rent

	coefficient	std. error	t-ratio	p-value	
const	51.2551	2.47257	20.73	0.0002	***
Price_to_rent_1	−0.538219	0.0236192	−22.79	3.80e-51	***
d_Price_to_rent_1	−0.0198787	0.0339739	−0.5851	0.5996	
d_Price_to_rent_2	−0.196550	0.0321261	−6.118	0.0088	***
d_Price_to_rent_3	−0.159702	0.0298835	−5.344	0.0128	**
d_Price_to_rent_4	−0.0269872	0.0253587	−1.064	0.3653	

AIC: 14.7245 BIC: 15.9078 HQC: 12.1708

Appendix 8. Testing for the presence of unit root in the time series for total homeownership rate

k = 4: AIC = 11.6370
k = 3: AIC = 13.8004
k = 2: AIC = 15.0531
k = 1: AIC = 13.3139
k = 0: AIC = 12.5483

test with constant
including 4 lags of (1-L)HOratetotalHH
model: (1-L)y = b0 + (a-1)*y(-1) + ... + e
estimated value of (a - 1): -0.97167
test statistic: tau_c(1) = -3.55311
asymptotic p-value 0.00675
1st-order autocorrelation coeff. for e: -0.467
lagged differences: F(4, 3) = 1.269 [0.4399]

Augmented Dickey-Fuller regression
OLS, using observations 2011-2019 (T = 9)
Dependent variable: d_HOratetotalHH

	coefficient	std. error	t-ratio	p-value	
const	79.7800	22.5188	3.543	0.0383	**
HOratetotalHH_1	−0.971670	0.273470	−3.553	0.0068	***
d_HOratetotalHH_1	−0.0117688	0.311383	−0.03780	0.9722	
d_HOratetotalHH_2	−0.237819	0.335060	−0.7098	0.5290	
d_HOratetotalHH_3	0.433742	0.225864	1.920	0.1506	
d_HOratetotalHH_4	0.208143	0.156691	1.328	0.2761	

AIC: 11.637 BIC: 12.8204 HQC: 9.08337

Appendix 9. Testing for the presence of unit root in the time series of the first difference of homeownership rates among households consisting of a single person

k = 4: AIC = 19.7539
k = 3: AIC = 18.3459
k = 2: AIC = 16.5649
k = 1: AIC = 14.5649
k = 0: AIC = 14.3698

Augmented Dickey-Fuller test for d_HOrateHHsingle
testing down from 4 lags, criterion AIC
sample size 12
unit-root null hypothesis: a = 1

test without constant
including 0 lags of (1-L)d_HOrateHHsingle
model: (1-L)y = (a-1)*y(-1) + e
estimated value of (a - 1): -1.23529
test statistic: tau_nc(1) = -4.06925
p-value 0.0007178
1st-order autocorrelation coeff. for e: 0.160

Dickey-Fuller regression
OLS, using observations 2008-2019 (T = 12)
Dependent variable: d_d_HOrateHHsingle

	coefficient	std. error	t-ratio	p-value	
d_HOrateHHsing~_1	−1.23529	0.303568	−4.069	0.0007	***

AIC: 22.7613 BIC: 23.2462 HQC: 22.5817

Appendix 10. Testing for the presence of unit root in the time series for homeownership rates among households consisting of two adults

k = 4: AIC = 10.9308
k = 3: AIC = 9.47231
k = 2: AIC = 9.17090
k = 1: AIC = 7.17092
k = 0: AIC = 9.83814

test with constant
including one lag of (1-L)HOrateHH2adults
model: (1-L)y = b0 + (a-1)*y(-1) + ... + e
estimated value of (a - 1): -1.21676
test statistic: tau_c(1) = -4.92192
asymptotic p-value 2.941e-05
1st-order autocorrelation coeff. for e: -0.157

Augmented Dickey-Fuller regression
OLS, using observations 2008-2019 (T = 12)
Dependent variable: d_HOrateHH2adults

	coefficient	std. error	t-ratio	p-value	
const	23.8970	4.85260	4.925	0.0008	***
HOrateHH2adults_1	−1.21676	0.247212	−4.922	2.94e-05	***
d_HOrateHH2adu~_1	0.514900	0.170014	3.029	0.0143	**

AIC: 6.754 BIC: 8.20872 HQC: 6.21541

Appendix 11. Testing for the presence of unit root in the time series for homeownership rates among households consisting of two adults and two dependent children

k = 4: AIC = 6.90636
k = 3: AIC = 18.2391
k = 2: AIC = 19.6759
k = 1: AIC = 18.6702
k = 0: AIC = 17.1584

test with constant
including 4 lags of (1-L)HOrateHH2adults2depchi
model: (1-L)y = b0 + (a-1)*y(-1) + ... + e
estimated value of (a - 1): -3.77005
test statistic: tau_c(1) = -4.85196
asymptotic p-value 4.061e-05
1st-order autocorrelation coeff. for e: -0.769
lagged differences: F(4, 3) = 4.949 [0.1099]

Augmented Dickey-Fuller regression
OLS, using observations 2011-2019 (T = 9)
Dependent variable: d_HOrateHH2adults2depchi

	coefficient	std. error	t-ratio	p-value	
const	49.4969	10.1936	4.856	0.0167	**
HOrateHH2adult~_1	−3.77005	0.777017	−4.852	4.06e-05	***
d_HOrateHH2adu~_1	2.49065	0.598530	4.161	0.0252	**
d_HOrateHH2adu~_2	2.00452	0.492469	4.070	0.0268	**
d_HOrateHH2adu~_3	1.56099	0.386811	4.036	0.0274	**
d_HOrateHH2adu~_4	1.16836	0.365873	3.193	0.0496	**

AIC: 6.90636 BIC: 8.08971 HQC: 4.3527

Appendix 12. Testing for the presence of unit root in the time series for homeownership rates among households with income over 60% of national equivalised net income

k = 4: AIC = 25.7761
k = 3: AIC = 23.7830
k = 2: AIC = 21.7890
k = 1: AIC = 21.9987
k = 0: AIC = 20.0229

Augmented Dickey-Fuller test for d_Owner_above_60_income
testing down from 4 lags, criterion AIC
sample size 12
unit-root null hypothesis: a = 1

test without constant
including 0 lags of (1-L)d_Owner_above_60_income
model: (1-L)y = (a-1)*y(-1) + e
estimated value of (a - 1): -0.692435
test statistic: tau_nc(1) = -2.4314
p-value 0.02009
1st-order autocorrelation coeff. for e: 0.061

Dickey-Fuller regression
OLS, using observations 2008-2019 (T = 12)
Dependent variable: d_d_Owner_above_60_income

	coefficient	std. error	t-ratio	p-value	
d_Owner_above_~_1	−0.692435	0.284789	−2.431	0.0201	**

AIC: 39.6238 BIC: 40.1087 HQC: 39.4443

Appendix 13. Testing for the presence of unit root in the time series for homeownership rates among households with income below 60% of national equivalised net income

k = 3: AIC = 33.0188
k = 2: AIC = 36.7209
k = 1: AIC = 35.1124
k = 0: AIC = 37.4677

Augmented Dickey-Fuller test for d_d_Owner_below_60_income
testing down from 3 lags, criterion AIC
sample size 8
unit-root null hypothesis: a = 1

test without constant
including 3 lags of (1-L)d_d_Owner_below_60_income
model: (1-L)y = (a-1)*y(-1) + ... + e
estimated value of (a - 1): -2.38685
test statistic: tau_nc(1) = -2.73939
asymptotic p-value 0.005978
1st-order autocorrelation coeff. for e: 0.028
lagged differences: F(3, 4) = 3.589 [0.1244]

Augmented Dickey-Fuller regression
OLS, using observations 2012-2019 (T = 8)
Dependent variable: d_d_d_Owner_below_60_income

	coefficient	std. error	t-ratio	p-value	
d_d_Owner_belo~_1	−2.38685	0.871307	−2.739	0.0060	***
d_d_d_Owner_be~_1	1.65553	0.743554	2.227	0.0900	*
d_d_d_Owner_be~_2	0.543510	0.435967	1.247	0.2805	
d_d_d_Owner_be~_3	0.544496	0.267009	2.039	0.1110	

AIC: 33.0188 BIC: 33.3365 HQC: 30.8756

Appendix 14. Initial regression model report (Model 1)

```
Model 6: OLS, using observations 2008-2019 (T = 12)
Dependent variable: own_total

                  coefficient    std. error    t-ratio   p-value
  ------------------------------------------------------------------
  const           74.4188        2.88750       25.77     3.38e-08 ***
  INT              1.50788       0.383807       3.929    0.0057   ***
  dd_income       -0.000564382   0.000492344   -1.146    0.2893
  d_PRICE         -0.0752401     0.0139189     -5.406    0.0010   ***
  Price_to_rent    0.0421927     0.0364422      1.158    0.2849

Mean dependent var   83.21667   S.D. dependent var   2.564738
Sum squared resid    1.319550   S.E. of regression   0.434174
R-squared            0.981763   Adjusted R-squared   0.971342
F(4, 7)              94.21015   P-value(F)           3.63e-06
Log-likelihood      -3.781565   Akaike criterion     17.56313
Schwarz criterion    19.98766   Hannan-Quinn         16.66548
rho                 -0.357518   Durbin-Watson        2.673719

Excluding the constant, p-value was highest for variable 15 (dd_income)
```

Appendix 15. Correlation matrix for the initial variables

Price_to _rent	own_total	INT	d_PRICE	dd_income	
1.000	-0.07451	0.7455	-0.3442	0.5271	Price_to_rent
–	1.000	-0.9154	-0.8287	-0.3023	own_total
–	–	1.000	-0.6213	-0.0717	INT
–	–	–	1.000	0.2021	d_PRICE
–	–	–	–	1.000	dd_income

Appendix 16. Adjusted regression model report with relevant tests (Model 2)

```
Model 2: OLS, using observations 2008-2019 (T = 12)
Dependent variable: own_total

              coefficient   std. error    t-ratio   p-value
  ------------------------------------------------------------
  const       77.7049       0.542093      143.3     6.27e-15 ***
  INT          1.89693      0.189362       10.02    8.38e-06 ***
  dd_income   -0.00101765   0.000304846    -3.338   0.0103   **
  d_PRICE     -0.0665325    0.0119587      -5.564   0.0005   ***

Mean dependent var   83.21667   S.D. dependent var   2.564738
Sum squared resid    1.572243   S.E. of regression   0.443317
R-squared            0.978271   Adjusted R-squared   0.970123
F(3, 8)              120.0568   P-value(F)           5.44e-07
Log-likelihood      -4.832843   Akaike criterion     17.66569
Schwarz criterion    19.60531   Hannan-Quinn         16.94757
rho                 -0.643731   Durbin-Watson        3.000806

Test for normality of residual -
  Null hypothesis: error is normally distributed
  Test statistic: Chi-square(2) = 2.36877
  with p-value = 0.305935

White's test for heteroskedasticity -
  Null hypothesis: heteroskedasticity not present
  Test statistic: LM = 11.8806
  with p-value = P(Chi-square(9) > 11.8806) = 0.220123

LM test for autocorrelation up to order 6 -
  Null hypothesis: no autocorrelation
  Test statistic: LMF = 0.93824
  with p-value = P(F(6, 2) > 0.93824) = 0.598285

LM test for autocorrelation up to order 1 -
  Null hypothesis: no autocorrelation
  Test statistic: LMF = 4.09669
  with p-value = P(F(1, 7) > 4.09669) = 0.0826379

RESET test for specification -
  Null hypothesis: specification is adequate
  Test statistic: F(2, 6) = 1.21419
  with p-value = P(F(2, 6) > 1.21419) = 0.360761
```

Appendix 17. Correlation matrix for variables used to model homeownership for low income group households

dd_low_inc_ owner	d_price	dd_low_inco me	Price_to_rent	INT	
1.000	0.3153	0.1466	0.27	0.3508	dd_low_inc_ owner
–	1.000	0.0526	-0.34442	-0.6213	d_price
–	–	1.000	-0.5589	-0.0970	dd_low_inco me
–	–	–	1.000	0.7455	Price_to_rent
–	–	–	–	1.000	INT

Appendix 18. Report for Model 3 with relevant tests

```
Model 13: OLS, using observations 2008-2019 (T = 12)
Dependent variable: dd_low_inc_owner

             coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const      -10.3504      3.32248      -3.115    0.0124  **
  INT          3.77149     1.15957       3.252    0.0100  ***
  d_PRICE      0.228112    0.0719039     3.172    0.0113  **

Mean dependent var   0.000000   S.D. dependent var   3.825750
Sum squared resid    66.65362   S.E. of regression   2.721389
R-squared            0.586002   Adjusted R-squared   0.494003
F(2, 9)              6.369628   P-value(F)           0.018901
Log-likelihood      -27.31488   Akaike criterion     60.62976
Schwarz criterion    62.08448   Hannan-Quinn         60.09117
rho                  0.096571   Durbin-Watson        1.622301

Test for normality of residual -
  Null hypothesis: error is normally distributed
  Test statistic: Chi-square(2) = 1.02337
  with p-value = 0.599485

White's test for heteroskedasticity -
  Null hypothesis: heteroskedasticity not present
  Test statistic: LM = 2.83847
  with p-value = P(Chi-square(5) > 2.83847) = 0.724872

LM test for autocorrelation up to order 1 -
  Null hypothesis: no autocorrelation
  Test statistic: LMF = 0.0741661
  with p-value = P(F(1, 9) > 0.0741661) = 0.79151

RESET test for specification -
  Null hypothesis: specification is adequate
  Test statistic: F(2, 7) = 4.74532
  with p-value = P(F(2, 7) > 4.74532) = 0.0498325
```

Appendix 19. Correlation matrix for variables used to model homeownership for higher income group households

INT	Price_to_rent	d_price	d_high_inc_owner	dd_high_income	
1.000	0.7455	-0.6213	0.0930	-0.1110	INT
–	1.000	-0.3442	-0.2422	-0.5098	Price_to_rent
–	–	1.000	0.0510	0.1631	d_PRICE
–	–	–	1.000	0.2131	d_high_inc_owner
–	–	–	–	1.000	dd_high_income

Appendix 20. Lihtlitsents

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks[1]

Mina, Kadri-Ann Freiberg

1. annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose "Determinants of Homeownership Rate in Estonia", mille juhendaja on Simona Ferraro,

1.1 reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh TalTechi raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2 üldsusele kättesaadavaks tegemiseks TalTechi veebikeskkonna kaudu, sealhulgas TalTechi raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. Olen teadlik, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autorile.

3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

04.01.2020.
/allkirjastatud digitaalselt/

[1]*Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil.*