The (in)equality of Housing Affordability within the Netherlands

Shawn R. Kuijpers (S3219682)

Abstract:

This paper examines the relevance of the consumption-adjusted housing affordability approach, developed by Ben-Shahar et al. (2019), to unravel potential growing(declining) (in)equalities and examine the potential elevation of housing affordability burdens among minority and underprivileged groups, with the use of the WoON datasets of 2009 up to 2018. The consumption-adjusted measure is replicated in a different context, (1) during a downward economic trend instead of an upwards trend and (2) within the Netherlands opposed to Israel in the original study. Findings suggest an overall decrease of inequalities between the more(less) burdened households within the Netherlands. However, simultaneously results suggest a trend of increasing inequalities between the lowest and highest income quantiles of the sample. Moreover, findings suggest noticeable increases of housing affordability burdens among underprivileged and minority groups. The consumption-adjusted measure provides, in line with the study of Ben-Shahar et al. (2019), an improved measure compared to the traditional price-to-income measure, which enables policymakers to mitigate affordability issues more accurately.

Keywords: Housing affordability, Price-to-income, Inequality

Colofon

Document	Master's Thesis Real Estate Studies
Title	The (in)equality of Housing Affordability within the Netherlands
Author	Shawn Rolf Kuijpers
Student number	S3219682
E-mail	shawn.kuijpers@gmail.com
Supervisor	Dr. M. (Mark) van Duijn
Assessor	S.L. (Sarah) Mawhorter, PhD
Date	August 12, 2022

Disclaimer: *"Master theses are preliminary materials to stimulate discussion and critical comment. The analysis and conclusions set forth are those of the author and do not indicate concurrence by the supervisor or research staff."*

Contents

1. INTRODUCTION

Rising rents and housing prices are a major topic of discourse in Europe. A shortage of affordable housing units is an increasing problem in Europe. Housing prices and rents constantly increased in comparison to incomes. Early 2021, the European parliament adopted a resolution in which its members are evoked to acknowledge fitting housing as a fundamental right that can be vindicated by legislation (Europees Parlement, 2021). Housing affordability issues may affect a household's budget, limiting the residual income left over for goods and services, and reduces the availability to save earnings. In general, housing affordability issues may result in decreased possibilities and a lower quality of life (Sawhill, 2018). Additionally, the availability of good quality, and affordable homes is declining (Ben-Shahar et al., 2019). Between 2010 and 2018 about 10 percent of the population in the EU spent more than 40 percent of their disposable income on housing. However, great differences exist between countries. Affordability issues do not only exist within Europe, but are a worldwide phenomenon. In Canada the CMHC indicates an affordability crisis due to a shortage of housing supply (CBC, 2022). In California lawmakers propose an assistance program to support first-time buyers to break the cycle of renting, whilst simultaneously suggesting a shortage of housing (KSBY, 2022).

Within the Dutch context, housing prices increased by 10% in 2020, being among the fastest growing markets within Europe (Calcasa, 2020). For the rental market, prices increased by 2,6% in the Netherlands in 2020, being the greatest increase of rent since 2014. On average, rents have increased by 2,63% yearly over the period 2009 - 2018, whilst inflation increased by 1,56% on average. For regulated rents the maximum increase of rents is determined by the inflation of the previous year and additional income dependent allowances. Another remarkable development is the stronger increase within highly urbanized areas, and Rotterdam exceeding Amsterdam in rent increases in 2020, averaging 2,8% for *blijvende huurders* and 4,1% for *bewonerswisseling* (CBS, 2020). For the private rental sector the government has also applied a maximum rent increase from 1 may 2021 onwards for three years, being 1% plus the inflation percentage, being 1,4% in 2020 (Rijksoverheid, 2021). This is the first time in recent history that the government intervenes in the Dutch private rental sector, as it used to be a 'free market'.

To counteract unaffordability, the Dutch government has taken measures for both renters and homeowners over the past years, such as the starters transfer tax reduction, in 2021, to encourage people up to the age of 35 to enter the housing market. For the rental market, a new form of housing allowance was introduced in 2006 by the Dutch tax authorities, as a means to compensate the population earning less than a 'modal income' (Nibud, 2021). The Netherlands ranks fifth of housing costs in disposable income in 2020, averaging 22,3% (Eurostat, 2022).

Homeowners also savour advantages on their mortgage, as the interest paid on loans are deductible for their income, resulting in tax reliefs. Legislation concerning the interest deductibility has changed over

the years however, and effects are limited over the next years, especially for homeowners who are within the highest tax bracket. This essentially means that disposable income of homeowners is positively affected by their mortgage in a sense compared to renters. Thus, for both homeowners and renters (within the social rent market) there are benefits which enhance affordability.

The housing stock in the Netherlands is divided as follows, about 58 percent of the tenants are homeowners and 42 percent renters in 2021 (CBS, 2021). The 42 percent renters can additionally be divided in two groups, the social rent-, and the private rental sector. The social rent sector is restricted and can never be higher than a certain amount as it is legally defined within the range of a point system and rent check. In 2022 a tenant is considered to be within the social rent sector when one pays below 763,47 Euros per month in basic rent and has to earn below a certain threshold in order to receive allowances. Once the threshold in number of points is exceeded, the house is considered to be within the private rental sector where no maximum rental price is legally determined. The Dutch government is considering to increase the threshold in legal terms to protect rents up to a price of a maximum of 1250 Euros per month by 2024, as rental prices are considered extortionate (NOS, 2022).

Appropriate policies are, however, important to be based on precise measures of affordability and financial distress. This paper attempts to address distress in the assessment of a normative quality- and consumption-adjusted measures of housing affordability based on the study of Ben-Shahar et al. (2019), and considers whether housing is affordable for both renters and homeowners. Renters are also considered within this study, in order to determine whether this group is able to afford housing, i.e. move from the rental market towards the homeowner sector.

1.1 Literature Review

Within the academic literature housing affordability is generally measured by relating housing costs to income, known as the house price-to-income ratio (e.g. Thalmann, 1999; Quigley and Raphael, 2004; Stone, 2006; Haffner and Boumeester, 2010). Traditional methods are continuously discussed whether they may be biased, as they do not account for variability in household compositions, consumption and preferences. For example households may prefer to consume excessive, or conversely, reduce housing services and be either regarded highly and less affordability burdened, respectively. Imagine a comparison between two parents and two children living in a two bedroom home and a single adult household with three bedrooms, comparing just income and housing expenses would provide a misleading perspective regarding housing affordability.

Within literature regarding housing (un)affordability, housing affordability is often portrayed as a strict percentage of income spent on housing (Bogdon and Can, 1997; Kutty, 2005). Over time the threshold determining housing affordability in the US shifted from 25 to 30 percent, taken from policies and housing

programs (Kutty, 2005). Thalmann (1999) both theoretically and empirically addressed the issue of merely using a strict percentage by measuring rent-to-income based on the average rent within the market. This thesis uses a novel method, which deals with the affordability issues by distinguishing apparent and actual affordability problems within the market.

Contemporary literature on housing affordability attempts to reduce the potentially biased overview of affordability by adjusting for consumption patterns, regional differences and prices indices over time (Gan and Hill, 2009; Ben-Shahar et al. 2019). As high cost-to-income ratios seem alarming, it might indicate different housing preferences for high quality and/or quantity housing (Lerman and Reeder, 1978). Therefore, policy makers should not be concerned with high ratios alone. Affordability metrics often underestimate concurrent housing affordability issues, as they do not recognize comprises households make in terms of housing characteristics in order to afford their homes. Larger households for instance, require more space and rooms in order to create a livable living space. Due to simplicity of the general price-to-income measure all households are generally examined together without considering the size, needs and opportunities of the households.

Within the Dutch context the stream of literature regarding housing affordability is dominated by research of Haffner and Boumeester (2010; 2014; 2015). Their research is primarily focused on traditional expenditure-to-income and rent-to-incomes measures of affordability within the Dutch context, limited to only present affordability ratios within a certain context (lowest 50 percent of incomes; energy costs, etc.). Whilst the research of Haffner and Boumeester sheds light on the affordability problems within the Netherlands, it does not however examine possible associations between different household compositions and their preferences.

This paper attempts to overcome this troubled view by using a novel normative affordability measure of Ben-Shahar et al. (2019). The normative measure used in this study corrects for the described bias in housing consumption. The measure is normative as it varies over space (peripheral levels), time (2009 – 2018), is derived endogenously, and it represents groups of similarly situated households, representing their consumption patterns. First, extensive datasets on Dutch households are used to bundle households based on geographic and household characteristics. Then, each household in the sample is matched to the average housing expense bundle of matching households. Thereafter a hedonic price index per region / province / periphery is estimated. Based on the computed data regarding normative household expenses, net income and pricing based on the hedonic pricing model, a consumption- and quality adjusted measure regarding housing affordability is computed.

This research paper focuses on the affordability and related affordability inequality within the Netherlands, and the differences between peripheral levels. The aim is to explore potential differences in affordability, with the use of a hedonic pricing model. Essentially, this paper, contrary to the current stream

of literature in the Dutch context, attempts to examine differences of households with the use of a consumption adjusted affordability measure to potentially unravel different causes for (un)affordability within its clusters.

1.3 Context – the period after the GFC

Due to the Global Financial Crisis (GFC) housing affordability has gained interest, as concerns grew related to the failing housing policies that lead to this crisis (Haffner and Hulse, 2021). The global crisis in 2008 was caused mainly by the collapse of the housing market, due to fragility of the mortgage market. Scholars tend to rethink policies constructed by the government. Li (2014) found that especially in the period of 2010-2011 a tremendous number of papers was published compared to the years before. The rapid growth in housing prices in the 00's of the twenty-first century caused scholars to investigate the ongoing growing gap between homeowners and renters, especially the 'generation rent', being the households who are entrants to the housing market, not able to buy their own property, and thus remain in the rental sector, often for the long term. This trend is of great interest, especially at current times, due to the rapid growth of housing prices in the EU, and the inability to enter the housing market as a 'starter'. Although hard to grasp on, price developments during the past years are similar to the growth in the years before the GFC in 2008. Wood et al. (2015) found that dynamic affordability issues were statistically higher for households in the years after the GFC, and the ability to sustain affordable housing became more difficult during these years.

Wetzstein (2017) denominates the trend of rapid rising housing costs compared to a slower rising wage and salary increase as the Global Urban Housing Affordability Crisis. In the worst cases households are forced to relocate to more rural areas due to the pressure of high housing costs in the dense urban areas. Younger generations, especially the lower middle class and the poorer are excluded from access to housing. In 2021 Wetzstein continued his work by the use of a qualitative study in five international cities to explore the decreasing housing affordability in urban areas. Although the accelerated housing expense rise faster compared to wage and salary increases is not unique, the housing affordability struggles occur simultaneously in urban areas across the world (Wetzstein, 2019; 2021). The fact that it's not merely a cyclical, but a structural issue pressuring housing affordability and the effect it has on more than half a billion households, supports the urgency of this topic both politically and academically to uncover the causes of the increasing affordability issues, and correctly respond with applicable policy changes.

In the twenty-first century the importance of the urban fringe gained more attention in relation to commuting costs as an additive to living costs. Urban markets are growing rapidly, and alongside unaffordability generally increases. Different measures are used to shed light to affordability dimensions in which house-price-to-income contributes tremendously. However, it is only one of the many used measures

which contribute to the literature highlighting a different dimension for instance compared a budget and/or ratio methodology of housing affordability. Commuting costs are not directly intertwined in the house-price-to-income method, however may affect the housing prices overall for both urban and rural areas (Ben-Shahar et al., 2020; Dewita et al., 2018). A mixture of dimensions related to housing affordability might help increase purposeful policy making.

1.4 Research Problem Statement

This paper attempts to fill the gap within the literature in two ways. First, this paper adds by examining an additional layer of affordability by measuring housing consumption per type of household based on household structure and their spatial context, secondly, the current stream of literature in the Netherlands regarding housing affordability limits itself to a general overview of the 'affordability crisis'. Household characteristics and household composition are important as housing preferences, comprises and needs differ per household composition and household characteristics. Cultural differences, such as ethnicity or country of origin might also affect household composition and size in relation to housing affordability. A larger family for instance, needs more rooms and space, and might have different preferences opposed to a single household. Housing (un)affordability can be linked to those preferences and needs of households, to measure whether these housing characteristics can be met. Additionally, preferences and quality of homes change of time, as Clark et al. (2000) researched over a period of 25 years.

In this research paper, the aftermath of the GFC is of importance, as the crisis was first visible within the housing market, households might experience comprises as what space to consume, the number of rooms needed, etc. Additionally, the spatial context is of interest, as different geographical locations allow comparison supporting identification of idiosyncrasies (Skaburskis, 1997). In the Netherlands preferences within different peripheral level might explain (un)affordability differences between urban and non-urban areas. As five levels of periphery are available within the dataset, the identification of potential idiosyncratic nature of each level can be compared. As stated by Burch (1996, p.64) "demographic factors come into operation … trends may well be influenced by local -even idiosyncratic- events, persons, or cultural factors.". and Skaburskis (1997 p.276) "The demand for housing units is affected by population size, by the manner in which the population divides into households and then by tenure choice, housing expenditures and locational preferences". This paper is an effects study, attempting to academically approach affordability by examining household characteristics. Following the method of Ben-Shahar et al. (2019) this paper provides a novel comparison to traditional measures of housing affordability within the Netherlands. Additionally, an attempt to reveal regional differences is considered, which might encourage more regional, contrary to national, policies. Therefore, this paper aims to answer the research question:

What is the relationship between household characteristics and locational characteristics and housing affordability in the Netherlands?

To answer the main research question, three sub-questions are formulated, distinguishing the theoretical and empirical paradigms, and essentially testing heterogeneity, respectively.

Sub-RQ1: What does theory explain about the concept of housing affordability?

To answer this sub-question, information from prior research has to be gathered and analyzed. Once the question is answered, a clear direction could be substantiated regarding the effect of household characteristics and locational characteristics on housing affordability. Theory will likely explain what factors account for the affordability of a household.

Sub-RQ2: What is the relationship between household composition, household characteristics and housing affordability after the GFC between 2009 and 2018?

To examine the effects of household composition, household characteristics and locational characteristics on housing affordability, a hedonic pricing model is applied. Earlier work by Ben-Shahar et al. (2019) sets a fundamental modelling basis on understanding the normative measure of housing affordability by stratification of groups within the dataset. Based on the research question, and prior literature, the method of Ben-Shahar et al. (2019) is adopted to create a conceptual model.

Sub-RQ3: What are the differences in housing affordability burdens between different levels of periphery?

To be able to answer sub question three, the WoON data is necessary to make a distinction in household composition, following the paper of Ben-Shahar et al. (2019). The dataset is stratified by clusters *ACPY,* being *A* Adults, *C* Children, *P* Periphery and *Y* Year. It is expected that effects will differ per level of periphery.

The remainder of this paper is constructed as follows. Section two further develops the theoretical basis in order to fully understand affordability. Section three describes the data and methodology in order to conduct the analysis. Section four presents the results from the conducted regressions. Finally, section five concludes the paper with a discussion, a summary and recommendations for further research.

2. THEORY

2.1 Measuring Housing Affordability

Housing affordability has become a widely used term in policy making and the public debate, however it remains a contested and poorly defined concept (Stone, 2006). A debate within the literature is to build consensus of the term 'affordability', a tension remains between the economical principle 'cannot pay' (Hancock, 1993) and the 'needs' principle (Whitehead, 1991).

The economical approach is generally interpreted and measured as the relationship between housing expenditures and household income. More specifically, affordability is measured by the ratio of housing costs and disposable income where measures of housing affordability differ at times, as norms change over time. As Hulchanski (1995, p. 2) states, "through the decades the housing expenditure-to-income 'rule of thumb' deemed to be an appropriate indicator of ability to pay gradually shifted upward… a 20 per cent rule lasted until the 1950s when somehow a 25 per cent rule came into use, only to be replaced in the 1980s by a 30 per cent 'rule of thumb'.", which remains a commonly used ratio. For instance, Rowley and Ong (2012) developed the 30/40 measure for Australian research, using 30 percent as a threshold for affordability and only examine the lowest 40 percent of incomes within the sample. Determined by a certain threshold, households are generally identified as either having an affordability problem or not.

The definition and principles used within this thesis are derived from Stone (2006), who advocates for a normative approach for housing affordability. The concept of housing affordability is "normative, requiring, to certain extent, an independent theoretical or logical foundation against which households' actual circumstances can be measured" (Stone, 2006, p. 157). Although using the definition of housing affordability of Stone (2006), who generally supports the residual income method, this methodology is not used within this paper, yet a normative approach is integrated within the 'needs principle' further explained below.

The needs principle is generally focused on the demand measure of particular groups in relation to the supply side, based on norms for these particular groups of households. Eventually determining what the norm of a household is depending on their resources and needs. Commonly used measures related to the needs principle are the mortgage-debt-to-housing price (e.g., Gyourko and Linneman, 1993; Norris and Shiels, 2007) and price-to-income measure (e.g., Kim and Cho, 2010; Gan and Hill, 2009; Zhang et al., 2016), in which the latter forms the basis for the dependent variable used within this paper. The price-to-income measure (PIR) is typically measured by the median ratio of the market price of a dwelling divided by the median income of a household in a geographical location, i.e. province or country. In this case renters are able to be included in the analysis, as the analysis examines whether housing units are affordable for the population as a whole, not simply the current homeowners. An example is the PIR measure used by

Galster and Lee (2021) showing a deterioration of housing affordability in Hong Kong, as the ratio developed from about 15 years of income in 2013 to almost 21 years of income in 2019. Housing units on general became more unaffordable for all countries examined in their research, indicating a housing affordability crisis. Galster and Lee (2021) "They report that the region's average price-to-income ratio is above 12,5, a staggering number indicating that housing is severely unaffordable for the large majority of urban dwellers. In particular, housing affordability tends to increase for cities with larger populations and lower household incomes in developing countries in Asia."

Evidence shows a fluctuation in the price-to-income over the years, often in line with the business cycles, as housing affordability in general increases during declines in housing prices, and conversely (Gan and Hill, 2009; Norris and Shiels, 2007; Ben-Shahar et al., 2019).

The basis of the price-to-income measure however, receives critique as Linneman and Megbolugbe (1992) summarized: 1) it underestimates burdens for low income households; 2) the measure does not control for the quality of housing stock in the market; 3) the price-to-income measure does not control for appreciation and changing housing costs, such as down payments and mortgage rates; 4) locational effects cannot be measured as it uses the median income of a certain group; and 5) it does not recognize high ratios related to changing household demands over the years.

The traditional price-to-income method developed over the years, which resolved some of the biases and critique delivered by Linneman and Megbolugbe (1992). Lerman and Reeder (1987) created a quality-based assessment of housing affordability issues based on housing costs that barely satisfy the adequacy criteria, introducing quality of housing in the housing affordability measure by the use of a normative measure for adequate housing based on minimal adequate rents standards of the Congressional Budget Office (CBO). Lerman and Reeder (1987) concluded that the conventional models overestimate housing affordability issues for renters and the quality-based housing affordability identifies different households as experiencing affordability issues in contrast to the conventional models. Bogdon and Can (1997) further deepen the quality-adjusted measurement of housing affordability by using the work of Lerman and Reeder (1987) and additionally adjusting for geographical distribution, comparing supply and demand in the Syracuse metropolitan area. Bogdon and Can (1997) examine affordability among the lower incomes, illustrating a potential mismatch in demand and supply of affordable homes for a specific region, essentially enabling policy makers to consolidate precise planning for areas of need. Fisher et al. (2009) continued the work of Bogdon and Can (1997) by adding a new dimension to locational effects, by incorporating amenity based affordability indexes, such as school quality, safety and job proximity. With the use of a hedonic pricing model for both renters and homeowners, they obtain implicit prices on locational effects as priorly described. The model developed by Fisher et al. (2009) sees affordability as an opportunity cost based on location, as they call it 'area affordability'.

The normative method of Ben-Shahar et al. (2019) addresses the critique on the traditional price-to-income method, and tackles the biases and critiques to a certain extent, as it includes and examines the burdens of low-income households, it partly corrects for the quality of houses, as it compares groups of similar households and their consumption, measures locational effects, and it addresses changing demands in housing characteristics, and thus demands over the years.

2.2 Theoretical Background of household characteristics (general overview)

The relationship between household characteristics and housing affordability is a substantial sub-theme within the housing affordability literature. Bujang et al. (2010), examine the relation between marital status, age, education and number of people within the household with housing affordability. Within their paper the price-to-income method is used with a threshold of 30 percent, concluding that household characteristics, such as education and marital status, show a considerable effect on housing affordability within the Malaysian context. Within the Canadian context Skaburskis (2004) examined housing affordability based on demography, geography, immigration, ethnicity, migration, income source, income recipients, education and employment. Education showed a marginal effect, changing job levels and income sources are key factors influencing housing affordability. Notably, affordability problems among young non-family households are growing. Immigration, ethnicity and migration result in a higher probability of housing affordability problems, however, are minor factors, and independent of the other factors. Within the Dutch context household characteristics are generally limited to age and household composition (Haffner and Boumeester 2010; 2014; 2015), excluding ethnicity in general.

However, most research is focused on the direct price-to-income measure in contrast to household characteristics. Ben-Shahar et al. (2019), introduced a consumption adjusted measure and thereafter examined the effects of household characteristics, concluding that household characteristics as ethnicity, education, marital status and age show effects on housing affordability within the Israelian context. The research of Ben-Shahar et al. (2019) shows a strong relationship between ethnicity and housing affordability, due to the household construction of neighborhoods and municipalities, sometimes living in segregated municipalities. Newman and Holupka (2014) examined housing affordability in relation to expenditure development of children, including few key predictors, such as ethnic background, education and locational effects. Similar to Ben-Shahar et al. (2019) Newman and Holupka stratified groups by individual, family and locational characteristics. The characteristics examined in Ben-Shahar et al. (2019) are also used within this thesis, each characteristic is discussed more in-depth, considering theory and international literature related to housing affordability.

2.3 Age (generation) and housing affordability

The association between age groups and housing affordability is widely examined (i.e., Mayer and Engelhardt, 1999; Wilcox, 2006; DeVaney et al., 2004). Wood et al. (2014) found that younger age groups, specifically under 35, experience elevated housing affordability burdens, opposed to other age groups within the Australian context. Unsurprisingly though, as younger adults, and first-time buyers, on average, do not earn as much compared to other age groups, and simultaneously have larger debts, i.e. study loans (Linneman and Megbolugbe, 1992).

Mayer and Engelhardt (1999) investigated change in homeownership in relation to age, concluding a decline, especially in younger adults up to 25 years and an increase in homeownership for relatively older households, especially for 65 and over. Likewise DeVaney et al. (2004) found a negative relation between age and housing affordability burdens, indicating that older age groups experience less affordability burdens compared to younger households.

Young households are considered to be between 20 and 39, which are also considered as the larger group of first-time buyers (Wilcox, 2006). The housing affordability issues experienced by these typical households are due to housing prices, income and choices of the household. The foremost reason of housing affordability issues is the limited housing supply (Zyed et al, 2016).

Within the Netherlands differences between age groups from 25 years and older are marginal, and the variable age does not have a strong relationship with affordability burdens according to Hoek et al. (2020). Their research found that the affordability ratio is higher for the age group between 25 and 34 compared to other age groups, although differences are marginal. Additionally, the housing cost ratio decreases as age increases until the age of retirement. Although younger households have a lower average income, their housing burdens are generally not extreme. This is due to the relatively lower quality homes and smaller spaces they consume. Younger age groups also choose more often to share housing, therefore their risk of experiencing affordability issues remains limited.

2.3 Education and housing affordability

There seems to be an association between education and housing affordability. Linneman and Megbolugbe (1992) claim that lower education households with middle-incomes are more prone to experience housing affordability issues. Turner and Luea (2009) found that education is a significant variable depicting housing demand and housing costs, related to housing affordability issues. Moreover, higher educated households are less prone to experience high housing affordability burdens compared to lesser educated households (DeVaney et al., 2004; Elmelech, 2004). DeVaney et al. (2004) found a positive relationship between education and housing affordability. Higher educated households are less likely to be severely housing affordability burden.

Similarly, Ben-Shahar et al. (2019) found that a one year increase of education relates to a 4,5 and 3,3 percent decrease in housing affordability burdens. Indicating a strong association between education and housing affordability. Note that the 3,3 (4,5) percent increase is related to the traditional (consumption-adjusted) measure, indicating an even stronger association for the consumption-adjusted measure.

2.4 Nationality, immigrant status and housing affordability

DeVaney et al. (2004) found that immigrants experience greater housing affordability burdens opposed to the American white household heads. Likewise, Murdie (2003) found that immigrants experience affordability issues in the rental sector in the city of Toronto, especially Somali immigrants. However, Murdie (2003) merely compares three groups of immigrants, not comparing it to 'native' and or later generation immigrants. Skaburskis (2004) also found that immigration, migration and ethnicity play a role in housing affordability issues within Canadian cities.

Ben-Shahar et al. (2019) show a similar association between nationality and housing affordability burdens. In Israel housing affordability burdens differ immensely between the base category, Jewish households, and Arab households. For the consumption adjusted measure almost 30 percent declines in housing burdens are found. Additionally, Ben-Shahar et al. (2019) examined the association between second generation nationality (i.e. father born in Europe, Asia or Africa) opposed to Israeli nationality as the reference variable. For these groups, only immigrants from the former Soviet Union experience significantly higher affordability burdens.

2.5 Marital status and housing affordability

Turner and Luea (2009) found an increase in housing demand and housing costs in relation to marital status. Marital status is also strongly associated with home ownership according to Hendershott et al. (2009), although no direct association between marital status and housing affordability is discussed, an indication suggests a higher possibility for couples to afford a home for themselves, instead of renting. Hendershott et al. (2009) also found a stronger association between wealth accumulation among couples compared to single households. However, separated couples experience wealth loss. As couples have combined wealth, in general housing affordability risks are limited, and couples are able to save at a faster rate compared to two singles, due to economies of scale. Note, that separation and divorce have a 'reverse effect' on housing affordability.

Ben-Shahar et al. (2019) found a strong association between marital status and housing affordability burdens, however did not find concluding differences between the consumption-adjusted- and traditional method. Therefore, a strong association is expected to be found between marital status and housing affordability burdens, without differences between the consumption-adjusted and the traditional method.

2.6 Hypotheses

Housing affordability in general has been extensively researched. The consumption-adjusted methodology however is a relatively new price-to-income methodology, which potentially unravels inequalities in consumption and quality of housing, especially for underprivileged households. As the method has been developed and examined by Ben-Shahar et al. (2019), it is important to investigate its strengths and weaknesses in other settings, i.e., countries and economic trends. First of all, the traditional price-to-net-income and the consumption-adjusted approach are compared and tested whether differences between the models are significant from one another, replicating the study of Ben-Shahar et al. (2019).

Thereafter, following prior research, and in order to test the usefulness of the model, the following hypotheses have been constructed based on the theoretical context:

I. Hypothesis I: age up to the age of retirement has a significant stronger negative association with housing affordability burdens compared to the age group 17-24.
II. Hypothesis II: when accounting for consumption patterns, people with less education are expected to have even greater housing affordability burdens as compared with people with more education.
III. Hypothesis III: nationality and immigrant status have a significant positive association with housing affordability burdens.
IV. Hypothesis IV: marital status is strongly related to housing affordability burdens.

In order to control for the development of (in)equalities over the years 2009-2018 within the Netherlands, a fifth, and final hypothesis is formulated as follows:

V. Hypothesis V: a decline in quality- and consumption-adjusted affordability is found.

3. DATA & METHODOLOGY

3.1 Data

This research paper examines data originated from the WoON survey, conducted by the Dutch ministry of BZK in cooperation with the CBS, additionally, the tax authorities supplemented data, i.e. data regarding income and subsidies. Since 2006 the survey replaced the WBO and KWR surveys, which are now combined. Prior to 2006 the survey had its focus on basic information regarding living. The WoON survey provides insight in housing, housing costs, household composition, housing requirements of households and the living environment. The research is conducted on national level, linked to location, on municipal and provincial levels, which causes data representativeness for the Netherlands as a whole. Note, that

although data is collected on regional levels, representativeness on regional level is limited. The raw datasets contain 67.071, 69.339, 62.668 and 67.523 responses, conducted in the years 2009, 2012, 2015 and 2018 respectively. Whilst this research was conducted, the 2021 WoON dataset has not been published yet. Therefore the 2018 is the most recent available data to this date.

In order to conduct the research certain observations are dropped and winsorized to remove outliers. First, if households include more than five adults, observations are dropped, due to the low amount of households (<0,14%), and the generalizability of the population. These groups are dropped as the number of clusters created, will extend by a large margin. Furthermore, observations are dropped if information regarding (1) WOZ-value, (2) Surface, (3) Building Age, (4) Type Home, (5) Tenant and (6) Ethnicity Generation is missing. Importantly, incomes below the legal standard of the 'Participatiewet', corresponding to the years, are dropped. As it accounts for the lowest allowance income. Below this legal standard incomes are so low, that it causes extreme affordability burdens. The triennial cross-sectional sample consists of a total of 218,137observations over the period 2009-2018, which remain after the removal of missing observations and dropping clusters containing less than 20 observations (described in the first step of the methodology), table 1 illustrates a compact overview of the samples examined, per year.

Table 1. Number of observations in the sample

Year	Raw Sample	Clean sample
2009	67,071	61,955
2012	69,339	54,931
2015	62,668	49,334
2018	67,523	51,917
Total	277,601	218,137

3.2 Methodology

The research conducted is based on the research of Ben-Shahar et al. (2019) who introduced a consumption-adjusted approach to measure housing affordability inequalities between households. However a major limitation and measurement issue arises within this study, as transactional data was not available within the dataset or in any other way possible to connect to the respondents of the WoON dataset. Therefore, WOZ value is used as a substitute variable, which is closest to property value[1]. WOZ value stands for 'Waardering Onroerende Zaken', which can be loosely translated to Valuation of Real Estate. Every year municipalities

[1] The correlation between recent transaction prices (2 year maximum) and WOZ value is 0,824, showing a strong relationship between the two variables.

determine the WOZ value of properties, which is based on an appraisal of similar real estate, compared to one another. When determining the WOZ value, the municipality uses an estimate of the value of a property being sold on 1 January the year before. Several issues arise when using this value, as (1) it is not a predictor of true market value of a property, (2) it lags behind one year and (3) it does not take quality of the real estate into account within the valuation. The WOZ value is however easily accessible, at least for all residential properties, and can function as an estimate of the property value.

Before analyses can be conducted several steps are required to prepare data in order to measure inequalities and variability of housing consumption. First the sample is stratified by year, locational and household characteristics. Specifically, the data is stratified by *A*, *C*, *P* and *Y*, which creates mutually exclusive clusters. Of which *A* denotes *Adults*, (1,2, 3 or more) per household. *C* denotes *Children*, (0,1,…,3 or more). *P* denotes *Periphery*, (1,2,…,5), being urban, outside urban center, green urban, village center and rural areas respectively. *Y* denotes *Year*, (2009, 2012, 2015 and 2018)[2]. In order to maintain representability per group, clusters containing less than 20 observations are deleted, following the methodology of Ben-Shahar et al. (2019).

Table 2 illustrates the households' number of people, denoted in children and adults. 34% of the households contain one adult, 64% comprises households with 2 adults and 1% of the households comprises of three adults or more. Within the households, approximately 65% of the households have no children, whilst 14% has 1 child, 15% has 2 children and 6% has 3 children or more. Substantial subgroups are couples without children and single households without children.

Table 2. Tabulation of children and adults within households in percentages related to the sample. N = 218,137

Children	Number of Adults			
	1 Adult	2 Adults	3 Adults or more	Total
No Children	28.23	35.49	0.93	64.65
1 Child	3.50	10.33	0.14	13.97
2 Children	1.89	13.38	0.08	15.35
3 Children or more	0.38	4.68	0.96	6.03
Total	34.00	63.89	2.11	100.00

The second step to recreate the consumption- and quality adjusted approach requires to create standardized housing consumption of household *i* in contrast to the stratum (ACPY) household *i* is part of. Housing consumption in this research is measured in the total surface household *i* consumes, which can be translated

[2] Example: ACPY = (2,2,2,2018) indicates that household *i* belongs to a specific cluster of the sample including 2 adults, 2 children, living in periphery 2, and is observed in the year 2018.

to living space in square meters[3]. Consumption is standardized by taking the total surface consumed in strata (i∈ACPY) divided by the number of households in strata ACPY. Therefore, the equation is as follows:

$$SC_{i\in ACPY}^{CA} = \sum_i SC_{i\in ACPY} / N_{ACPY} \qquad \text{eq. (1)}$$

Where $SC_{i\in ACPY}$ denotes the total surface consumed by household i in stratum ACPY. N_{ACPY} denotes the number of households in stratum ACPY. $SC_{i\in ACPY}^{CA}$ essentially denotes the average surface consumed for a household within a certain cluster. As Ben-Shahar et al. (2019) states "by grouping households by ACLY and accordingly deriving ($NR_{i\in ACLY}^{CA}$), we essentially eliminate the inherent positive correlation between i's income and housing consumption." Positive correlation is eliminated by specifically deriving the housing consumption of a typical set of household characteristics within a comparable peripheral area within a cluster.

Hereafter, the dataset is divided by the 12 provinces of the Netherlands, creating 12 separate datasets, with observations ranging between 3,927 (Drenthe) up to 64,190 (Zuid-Holland). This step is not performed within eq. (1), due to the substantial decrease of observations if provinces were included in the stratification. Note that municipal or city levels are not available within the dataset, therefore provinces were used to provide a more specific geographical specification to examine compared to the sole use of periphery. Although representativeness of regions is limited due to missing weighted averages within the WoON dataset for regional level, provinces can provide a more relatable geographical environment compared to peripheral levels. A hedonic pricing model is constructed to compute estimated housing prices (WOZ-values) per province. The hedonic pricing model for object *o* in province *p* is constructed as follows:

$$\ln(HP_{op}) = \beta_0 + \beta_1 S_o + \beta_2 C_o + \beta_3 TFE_o + \varepsilon_{op} \qquad \text{eq. (2)}$$

Within this model, the y-intercept is denoted by β_0, the coefficients are noted as, β_1 , β_2 and β_3. The dependent variable *HP* denotes the WOZ-value at that given time. *S* denotes the surface of the object, *C* denotes the control variables related to property characteristics, being (1) building age, (2) stories and (3)

[3] In the Netherlands a commonly used measure within the housing market is 'woonoppervlakte', or surface. This is therefore used as an indicator of housing consumption. Other measures for have been tested as well, being (1) number of rooms and (2) surface categorized (seven groups). However, these models showed sharp declines in the explained variance of the independent variables on the dependent variable (R^2). Therefore, only the surface consumed has used to determine housing consumption.

DumNew, a dummy variable noting 1 as an objects' age is up to 2 years and otherwise 0[4]. TFE is the time fixed effect within this model by year, ln(x) is the log operator and ε denotes the error term of the equation. Note that the equation is computed separately for every province. Importantly, the coefficients are noted as they are essential to compute equation (3), later presented. Table 2 presents the outcome of the coefficient estimates for all provinces separately. All models are robust and significant at a 1-% level. R-squares of the computations are between 0,320 (Zeeland) and 0,511 (Zuid-Holland). After the computation of equation (2) for every province, the datasets are merged. The saved estimations of the coefficients are a preliminary step to constitute a consumption adjusted measure of affordability.

Following equation (2) a consumption-adjusted standardized housing price is computed. The hedonic pricing model for household *i* in province *p* is constructed as follows:

$$HP^{CA}_{i\in ACPY} = EXP[\hat{\beta}_0 + \hat{\beta}_1 SC^{CA}_{i\in ACPY} + \hat{\beta}_2 \overline{C} + \hat{\beta}_3 TFE_{i\in ACPY} + \hat{\varepsilon}^2_p/2] \text{ for all } i \text{ and } p \text{ - eq. (3)}$$

Within this model, the y-intercept is denoted by the estimated coefficient $\hat{\beta}_0$ originated from equation (2), the estimated coefficients of equation (2) are noted as, $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$. The dependent variable HP denotes the consumption adjusted WOZ-value at that given time. $SC^{CA}_{i\in ACPY}$ denotes the standardizes surface consumed by household *i*, $\overline{C}$ denotes the averages of characteristics (building age and stories) within cluster $i\in ACPY$. TFE is the time fixed effect within this model by year, ln(x) is the log operator and ε denotes the estimated error term from equation (2). A hedonic price, $HP^{CA}_{i\in ACPY}$ is computed for each household *i* in cluster $i\in ACPY$, corresponding with its consumption-adjusted bundle, $SC^{CA}_{i\in ACPY}$.

Additionally, equation (3) is transformed to compute the actual housing consumption of household *i*, in order to compare the two approaches. The hedonic pricing model for the traditional approach is constructed as follows: $\hat{HP}_{i\in ACPY} = EXP[\hat{\beta}_0 + \hat{\beta}_1 SC_{i\in ACPY} + \hat{\beta}_2 \overline{C} + \hat{\beta}_3 TFE_{i\in ACPY} + \hat{\varepsilon}^2_p/2]$ for all *i* and *p* - eq. (4)

Note, that $SC_{i\in ACPY}$ denotes the actual surface consumed by household *i* in cluster *ACPY* respectively. $\hat{HP}_{i\in ACPY}$ denotes the unadjusted WOZ-value of household *i*.

After the computation of the consumption-adjusted and traditional housing consumption, the housing affordability measure is constructed with household *i*'s *Income*. The housing affordability, i.e.,

[4] Both building age and DumNew are included as control variables, as building age gradually explains changes in WOZ value, whilst newly built properties (DumNew) show substantial increases in WOZ value compared to relatively older buildings.

price-to-net income ratio as explained in the theory section, is computed as follows: $\widehat{HP}_i^{CA}$ / $Income_i$ (consumption-adjusted price-to-net income ratio) and $\widehat{HP}_{i \in ACPY}$ / $Income_i$ (traditional price-to-net income ratio). This forms the basis to conduct analyses, presented in the following section and the results chapter.

3.3 Descriptive statistics

The descriptive statistics in table 3 and 4 provide a general overview of the variables used in the analysis. Differences between variables of interest are *Surface* and $Surface_{CA}$, where a distinction is made between the actual surface 'consumed' by household *i* and the consumption adjusted surface consumed by cluster *ACPY* respectively. Another key difference in variables is noticed in HP_{TR} and HP_{CA}, being the traditional- and consumption adjusted estimated WOZ-values, essential for the regressions, further explained in the results section.

Table 3. Descriptive statistics. N = 218,137

Variable	Definition	Obs	Mean	Std. Dev.	Min	Max
Age	Age of the respondent categorized	218,137	4.31	1.631	1	7
Adults	Number of adults in household	218,137	1.681	.509	1	3
Children	Number of children in household	218,137	.628	.950	0	3
Partner	Does the respondent have a partner (0 = no / 1 = yes)	218,137	.639	.480	0	1
Household composition	Type of household composition	.	.	.	.	.
One person household		218,137	.282	.450	0	1
Couple		218,137	.346	.476	0	1
Couple with children		218,137	.293	.455	0	1
One parent family		218,137	.060	.238	0	1
Non-family household		218,137	.018	.134	0	1
Ethnicity	Ethnicity of the respondent	.	.	.	.	.
Non-immigrant		218,137	.844	.363	0	1
First generation		218,137	.090	.286	0	1
Second generation		218,137	.066	.248	0	1
Country Born	Country of origin of the respondent	.	.	.	.	.
Netherlands		218,137	.900	.300	0	1
Non-western country		218,137	.059	.235	0	1
Western country		218,137	.041	.199	0	1
Education	Education level of respondent	.	.	.	.	.

Low educated		218,137	0.340	0.474	0	1
Middle educated		218,137	0.329	0.470	0	1
Highly educated		218,137	0.331	0.471	0	1
Province	All the provinces of the Netherlands	.	.	.	.	.
Groningen		218,137	.024	.152	0	1
Friesland		218,137	.023	.150	0	1
Drenthe		218,137	.018	.133	0	1
Overijssel		218,137	.065	.247	0	1
Flevoland		218,137	.027	.163	0	1
Gelderland		218,137	.150	.357	0	1
Utrecht		218,137	.058	.233	0	1
Noord-Holland		218,137	.120	.325	0	1
Zuid-Holland		218,137	.294	.456	0	1
Zeeland		218,137	.048	.215	0	1
Noord-Brabant		218,137	.118	.322	0	1
Limburg		218,137	.054	.226	0	1
Periphery	Periphery ranging from 1 (urban)	.	.	.	.	.
Urban center	to 5 (rural)	218,137	.060	.238	0	1
Urban outside center		218,137	.416	.493	0	1
Green Urban		218,137	.112	.316	0	1
Center Village		218,137	.313	.464	0	1
Rural		218,137	.099	.299	0	1
Tenant	Type of tenant	.	.	.	.	.
Social Rent		218,137	.279	.448	0	1
Private Rent		218,137	.077	.266	0	1
Homeowner		218,137	.644	.479	0	1
Income	Income per household	218,137	38705	21575	10878	127965
Surface	Actual surface consumed	218,137	119.6	61.019	35	398
Surface$_{CA}$	Consumption-adjusted surface	218,137	119.6	24.117	80.445	198.688
HP$_{TR}$	Traditional approach – WOZ value	218,137	223367	89178	18511	1024327
HP$_{CA}$	Consumption-adjusted WOZ value	218,137	217406	57365	15494	664927

Table 4. Descriptive statistics. N = 218,137

Variable	Mean	Std. Dev.	Min	Max
WOZ	235643	125915	67000	789000
Surface	119.6	61.019	35	398
Building Age	44.489	29.486	0	149
Stories	2.12	1.086	1	85
DumNew	.015	.12	0	1

4. RESULTS AND DISCUSSION

This chapter presents the results from the ordinary least square (OLS) models. Section 4.1 discusses differences between the traditional- and consumption adjusted approach by visualizing data on household characteristics, economic and geographic levels. Section 4.2 presents main model, based on both the traditional- and consumption adjusted approach. Within this regression the household characteristics are central as introduced in the theory section. Section 4.3 tests robustness of the model on national level with the use of economic wide trends, and simultaneously transforming the dependent variable to a relative-to-the-mean variable. Four models are presented, testing different groups and pools. Section 4.4 presents the model developed in section 4.3 distincting rural and urban areas to test the housing affordability burden differences between rural and urban areas within the Netherlands.

4.1 Visualization of differences between the traditional and consumption-adjusted approach

Within this chapter general visualization of the models are presented to form basis for the results of the ordinary least squares (OLS) regressions. Both the traditional- and consumption adjusted methods are transformed to graphs to visualize potential differences on geographic, specific household characteristics and economic levels. Comparisons are made between broader aspects such as country wide analyses and regional analyses and niche aspects, being comparisons between certain groups, such as tenure mode, education, nationality and gender. Further regressions of the full models are presented, and differences discussed, and at last the models of Equation 6 attempt to reveal economy wide affordability burdens between groups on both national and provincial levels.

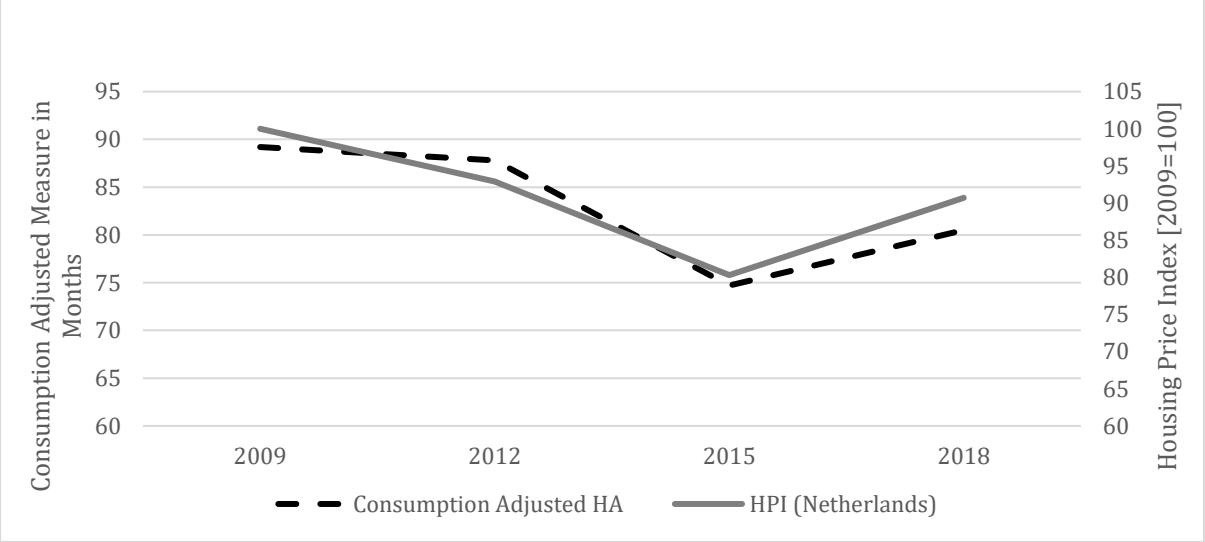


Figure 1. Consumption-adjusted housing affordability measured in months of residual income (HP_{CA} /Income) and the House Price Index (HPI) in the Netherlands between 2009 and 2018. HPI base (2009 = 100). N = 218,137

Figure (1) presents the alignment of the general housing affordability burden within the consumption adjusted method and the house price index (HPI) of the Netherlands[5]. The graph shows the course of development of the consumption-adjusted price-to-income ratio is closely related to the nationwide HPI. Correlation matrix shows a strong connection (0,93) between the two variables. Prominent is the increase in housing affordability between 2012 and 2015, during the measured years, the effects of the GFC were most pronounced in 2015, where the house price index was lowest, the affordability was highest, aligning predictions.

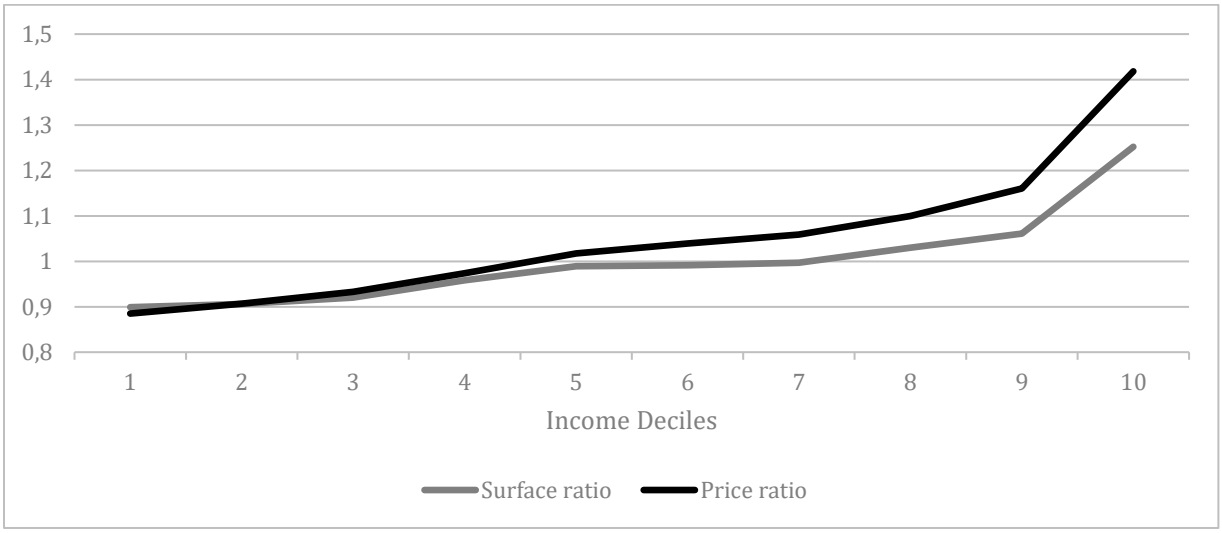


Figure 2. Surface ratio (Surface / $Surface_{CA}$) and the Price ratio (WOZ /WOZ_{CA}) per income decile in 2018. N = 51,917

[5] The WOZ-valuation dates are in most cases corresponding with a year prior to the dataset, thus n-1. The WOZ-values per year are as followed: 01-2009, 01-2011, 01-2014 and 01-2017 for the datasets 2009, 2012, 2015 and 2018 respectively. Note that the HPI is corresponding to the valuation dates.

Figure (2) illustrates the ratio of the surface area and house price respectively of household *i* to the average surface area and estimated house price of the stratum it's in (ACPY) for the year 2018. These ratios are measured in income deciles of the total population. Until the 8th decile the surface ratio is below 1, meaning that on average household *i* has a smaller surface compared to the stratum, for house price the 5th decile exceeds one. For both ratios a strong increase is visible between the 9th and 10th decile, meaning that inequalities are most present in the last Incomes deciles. The surface ratio lies between 0,899 and 1,252, a 0,353 difference is thus seen between the lowest and highest income decile respectively. For the ratio of house price, the ratio is 0,967 and 1,231 for the lowest and highest income decile respectively
. Indicating a higher degree of equality, 0,264, between house price compared to surface ratios.

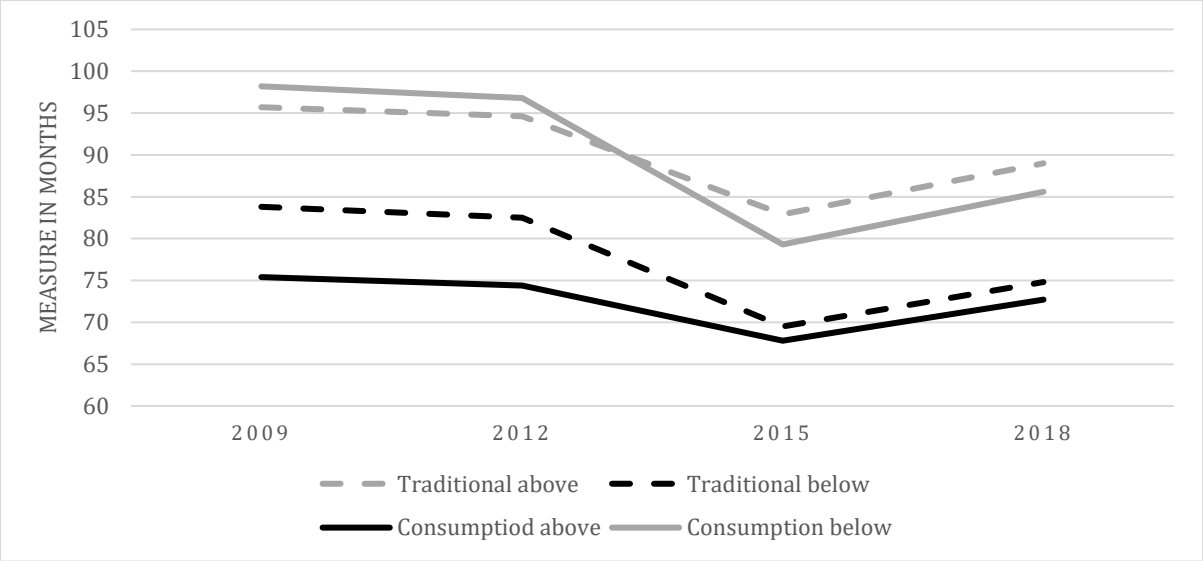


Figure 3. Traditional- and consumption-adjusted affordability measured in months of residual income (HP_{CA} and HP_{TR} /Income).. Above (below) are households who exhibit Surface ≥ $Surface_{CA}$ (Surface < $Surface_{CA}$) respectively. N = 218,137

Figure (3) illustrates the differences between the traditional and consumption adjusted model used within this paper. The traditional measure calculates the standardized house price with the use of the actual surface used by household *i* compared to the standardized surface used within the consumption adjusted method. Groups are stratified by consumption, groups who consume an excess(less) than their corresponding stratum (ACPY). This analysis provides a general insight in the housing affordability within the Netherlands. Additionally, a distinction is made between the two methods. Figure 3 indicates a greater difference between the above and below consumption adjusted groups between 2009 and 2012 respectively, whilst differences between these groups decrease in 2015 and 2018. The traditional measure develops in a more evenly manner.. The outliers between the consumption adjusted groups (above and below) show an average difference of 20,6 months and 10,2 months in 2009 and 2015 respectively. Whilst I averages of months between the two groups between 2009 and 2018 for the traditional measure lie between 14,2 and 15,4 (indicating a maximum change of 1,2 months between the groups over the years). Inequality is thus

more pronounced in the higher affordability burdened years (2009 – 2012) and less pronounces in the relatively lower affordability burdened years (2015 – 2018) for the consumption adjusted method. Note that the outcomes between the two methods of Figure 3 differ significantly at a 1%-level.

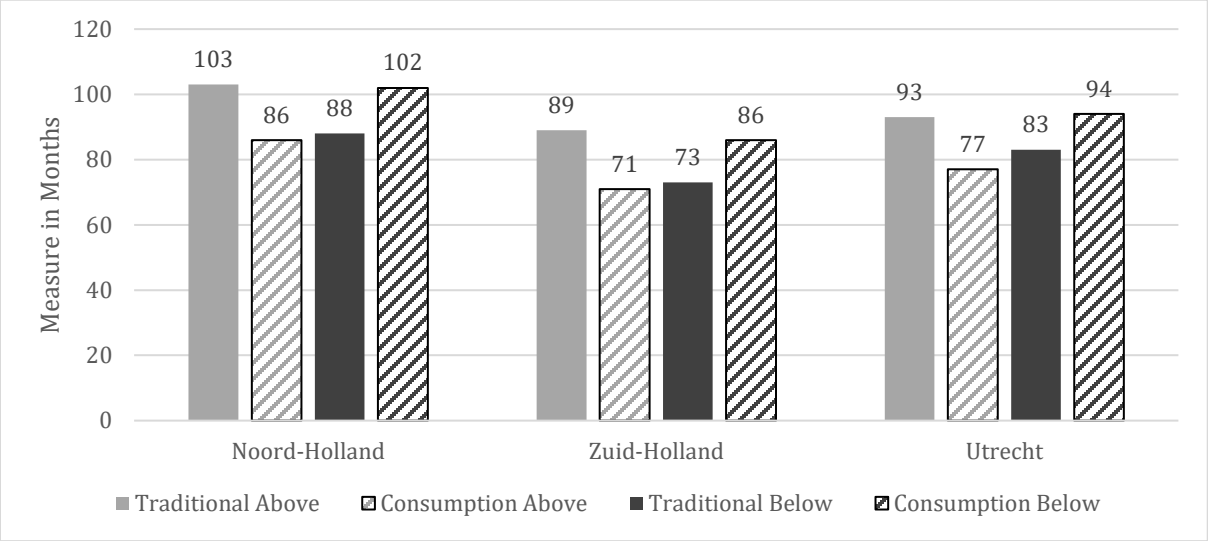


Figure 4. Traditional- and consumption-adjusted affordability measure in months of residual income (HP_{CA} and HP_{TR} /Income) on Provincial level within the Randstad.. Year 2018. N = 22,112

Figure (4) illustrates between group differences computed in Figure (3) within the Western region of the Netherlands, being the Randstad, per province in 2018. Groups are stratified by housing consumption above and below the consumption adjusted standard. All groups presented in Figure 4 for both below and above consumption adjusted groups differ significantly from one another at the 1%-level in t-tests. When comparing group differences, the traditional above and below differences averages 14 months in difference, whilst the differences in the consumption adjusted approach differ 16 months on average. Differences between the traditional- and consumption adjusted approach are on average 14,3 percent, indicating that the consumption adjusted measure examines higher levels of inequality within the Randstad. This approach is therefore able to provide insights regarding inequalities in housing affordability within the provinces[6]. Separately, these results are tested on three other regions in the Netherlands, being North (Drenthe, Groningen and Friesland), East (Overijssel, Gelderland and Flevoland) and the South (Zeeland, Noord-Brabant and Limburg), illustrated in Appendix (D). Differences between the traditional- and consumption adjusted approach are more apparent in these three regions with differences ranging between 30,0% (North) and 100,0% (South), indicating stronger inequalities in these regions compared to the West region (Randstad and/or Green Heart) of the Netherlands.

[6] Although weighted averages per households are available within the WoON dataset, these weights are not used within this paper, due to the complexity and the reliability of these weights, as they are criticized.

Furthermore, robustness is tested by stratifying household groups by a variety of characteristics, mainly household characteristics. Figure 5 to 8 illustrate the stratification per group, including tenure (renter or homeowner), nationality (immigrant or non-immigrant) and education level (low, middle or highly educated).

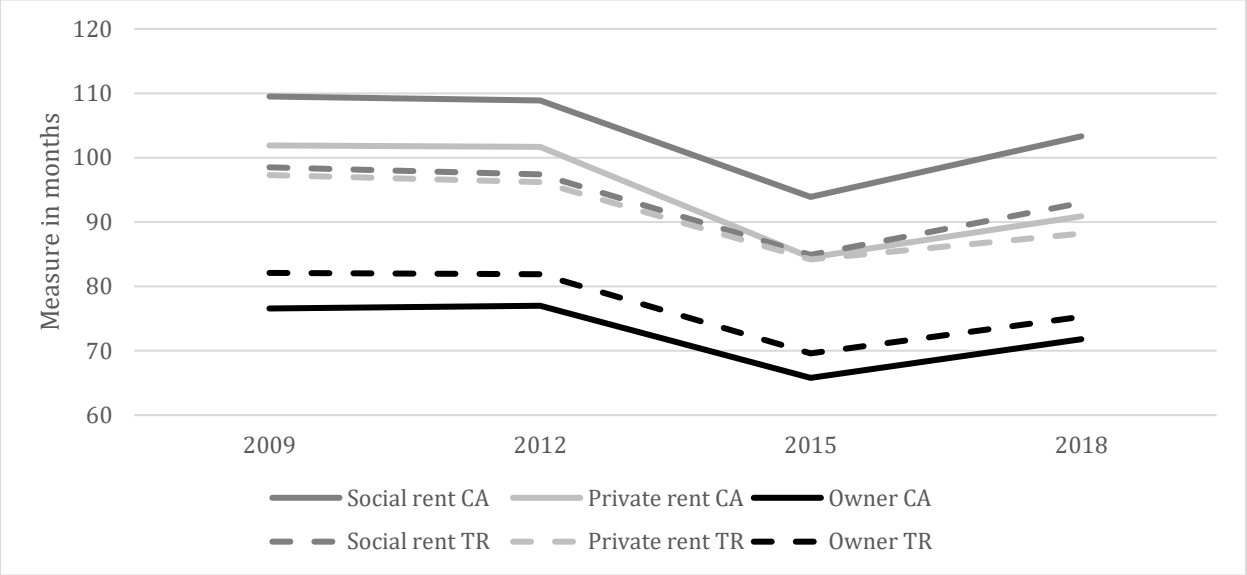


Figure 5. Traditional- and consumption-adjusted affordability measure in months of residual income (HP_{CA} and HP_{TR} /Income). Tenure, homeowner vs. social and private renter. N = 218,137

Figure 5 illustrates the in-between differences per tenure type. These results differ significantly from one another on a 1%-level. The average gap in months between renters and homeowners within the traditional- and consumption adjusted approach amount to 14 and 27 months respectively. Within the consumption adjusted methodology larger gaps are present compared to the traditional method indicating that homeowners experience less affordability burdens compared to the traditional method whilst renters experience larger housing affordability burdens compared to renters within the traditional method. This essentially indicates lesser burdens for homeowners (the group that is least burdened already) and indicates growing burdens for renters (which already experience relatively high burdens). The consumption adjusted method therefore recognizes burdens for the least affordable groups in a more profound way.

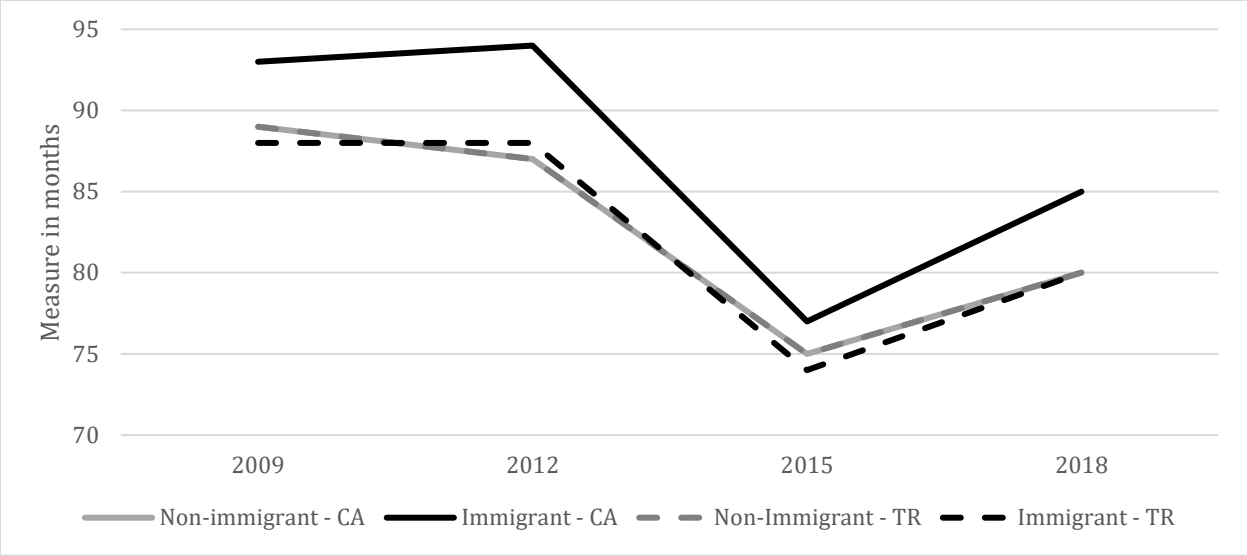


Figure 6. Immigrants vs. non-immigrants. Traditional- and consumption-adjusted approach. Measure in months. N = 218,137

Figure 6 recognizes in-between group differences for non-immigrants and immigrants within the Netherlands. The in-between group differences for the traditional method are on average -1,3 months for immigrants, indicating that immigrants experience less affordability burdens compared to non-immigrants, whilst the consumption adjusted approach indicates an in-between group difference of 4 months in favor of non-immigrants. This indicates that the consumption adjusted approach recognizes immigrants within the Netherlands as relatively more housing affordability burdened, which is in line with the literature (Ben-Shahar et al., 2019; DeVaney et al., 2004), forming a preliminary basis supporting hypothesis III. Although the sensitivity of the topic, in a logical sense the expectation is that immigrants experience a higher burden opposed to autochthonous residents. Therefore, it is unique to see the difference between the traditional- and consumption adjusted approach, as they contradict one another. Although differences between non-immigrant and immigrant residents within the traditional approach are marginal, the differences are significantly tested at a 1%-level.

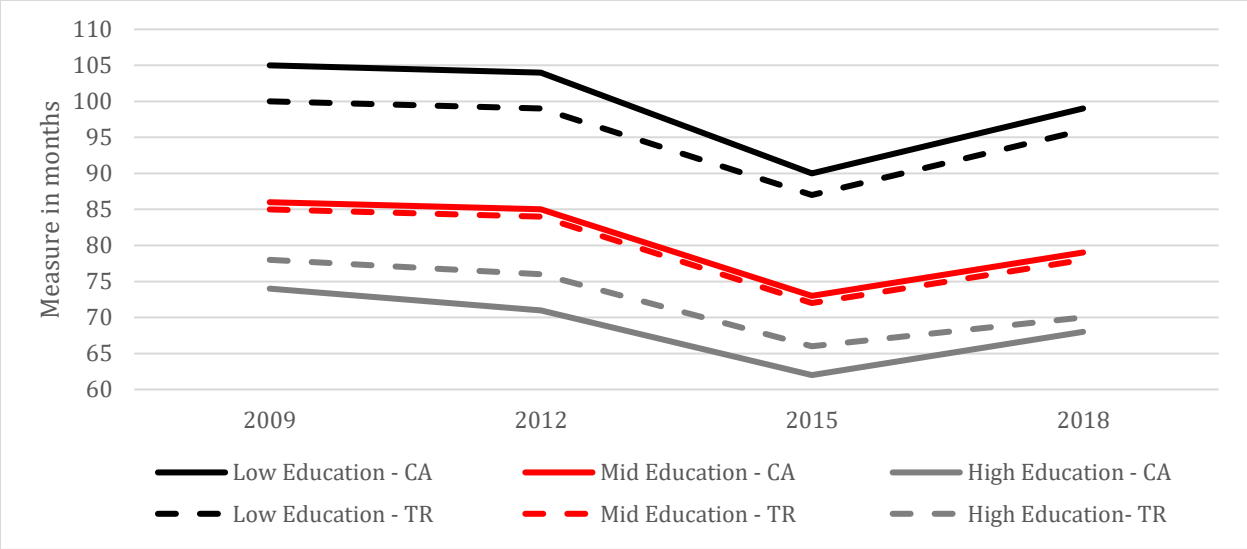


Figure 7. Education level of respondents corresponding to their housing affordability burdens. Traditional- vs. consumption-adjusted approach. Measure in months. N = 218,137

Figure 7 illustrates the differences between the traditional- and consumption adjusted approach by education, being either (1) low educated, (2) middle educated, and (3) highly educated. Similar to the outcomes of Figure 5 and 6, the consumption adjusted approach recognized a larger gap between the lesser and more privileged groups. Figure 7 indicates that the low(highly) educated residents experience relatively higher(lower) housing affordability burdens compared to the traditional approach. The gap between the low and highly educated groups within the consumption adjusted approach are on average 29,6 months, whilst the gap for the traditional approach averages 21,8 months. Yet again, the consumption adjusted approach indicates a more pronounced effect on highly and relatively low burdened households. For the highly burdened groups it shows an increase in burden, whilst for the least burdened group a reduction in burden is found.

4.2 The main model, traditional- and consumption-adjusted approach

This section examines the relation between household characteristics and housing affordability. The model is partly based on the study of Ben-Shahar et al. (2019). The model is as follows:

$$ln(HA_{iy}) = \beta_0 + \beta_1 Country_i + \beta_2 Immigrant_i + \beta_3 Age_i + \beta_4 Education_i + \beta_6 Composition_1 + \beta_7 Periphery_i + \beta_8 TFE_i + \beta_9 LFE_i + \varepsilon_{iy} \quad \text{eq. (5)}$$

In which HA_{iy} = **(**HP_{iy}^{CA} / $Income_{iy}$; HP_{iy} / $Income_{iy}$) represents the consumption-adjusted- and traditional measure respectively. In the equation i denotes household and y denotes the year of the measure. First the β_0 denotes the intercept of the dependent variable *HA*, coefficients of the independent variables are

indicated by β_1, β_2 … β_9. *Country* denotes the country of origin of the respondent, *Immigrant* denotes the ethnicity of the respondent, *Age* is the age of the respondent, *Education* denotes the highest education level achieved by the respondent, *Composition* denotes the household composition, *Periphery* denotes the level of periphery the respondent lives in, *TFE* and *LFE* indicate the time-fixed- and locational-fixed effects respectively and ε is the error term in the equation. This equation enables to distinct differences between the traditional- and consumption model, which are presented in table 5.

Table 5. Outcomes obtained from the regressions based on eq. (5), the consumption-adjusted and traditional measure.

VARIABLES	(1) Consumption-adjusted	(2) Traditional	(3) Differences in coefficients
Constant	2.253***	2.103***	
	(0.00828)	(0.00838)	
Country Born			
Dutch	Base	Base	
Non-western	0.0786***	0.0639***	
	(0.00952)	(0.00921)	
Western	-0.0117	0.0172*	***
	(0.00902)	(0.00879)	
Ethnicity			
Non-immigrant	Base	Base	
First generation immigrant	0.0572***	0.00405	***
	(0.00932)	(0.00905)	
Second generation immigrant	0.00792**	0.00512	
	(0.00321)	(0.00320)	
Age			
Age 17-24	Base	Base	
Age 25-34	-0.123***	-0.107***	
	(0.00583)	(0.00594)	
Age 34-44	-0.181***	-0.0972***	***
	(0.00588)	(0.00598)	
Age 45-54	-0.252***	-0.127***	***
	(0.00585)	(0.00595)	
Age 55-64	-0.194***	-0.0424***	***
	(0.00585)	(0.00596)	
Age 65-74	-0.0466***	0.103***	***
	(0.00589)	(0.00601)	
Age 75 and older	0.0133**	0.145***	***

	(0.00604)	(0.00618)	
Education			
Low educated	Base	Base	
Secondary educated	-0.126***	-0.0667***	***
	(0.00199)	(0.00201)	
Highly educated	-0.297***	-0.166***	***
	(0.00207)	(0.00207)	
Household composition			
One person household	Base	Base	
Couple	-0.341***	-0.349***	
	(0.00197)	(0.00199)	
Couple with children	-0.471***	-0.465***	
	(0.00229)	(0.00231)	
One parent family	-0.124***	-0.114***	
	(0.00400)	(0.00399)	
Non-family household	-0.283***	-0.286***	
	(0.00757)	(0.00782)	
Observations	218,137	218,137	
R-squared	0.404	0.372	
TFE	Yes	Yes	
LFE	Yes	Yes	

Note: dependent variables for model 1 and 2 are the consumption-adjusted- and traditional measure respectively ($HA_{iy}^{CA,}$; HA_{iy}). Robust standard errors in parentheses and significance depicted with *** p<0.01, ** p<0.05, * p<0.1 respectively.

At first glance, the traditional and consumption-adjusted models do not vary greatly. However differences in the models are highlighted in the third column indicating the most pronounced differences between the models, following the analysis of Ben-Shahar et al. (2019). The consumption-adjusted measure does significantly differ compared to the traditional measure. Furthermore, in order to answer hypothesis I – IV, each household characteristic variable is separately discussed in the following paragraphs.

Age is a significant determinant of housing affordability burdens within the consumption-adjusted measure. The base, or reference group are respondents between the age of 17 and 24. If the age category elevates, in general a decline of housing affordability burdens is seen. Especially for the age group 45-54 whose affordability burdens are 25,4% lower compared to the reference group. Note that age groups from 65 onwards experience relatively higher housing affordability, due to a lower average income, since most respondents are retired at that age. Outcomes are in line with prior literature (Mayer and Engelhardt, 1999; Devaney et al., 2004; Wood et al., 2014). Additionally, outcomes of the consumption-adjusted measure

suggest considerable declines in housing affordability burdens, in line with Ben-Shahar et al. (2019). Therefore, Hypothesis I is supported.

Also, education is a strong determinant of housing affordability burdens, especially within the consumption-adjusted measure. The base, or reference group are the respondents who achieved the low education levels. For the middle educated population a 12,6 percent decrease in housing affordability burdens is found, and for the highly educated population a 29,8 percent decrease in housing affordability burdens is found, indicating a strong negative association between education level and housing affordability burdens. Vice versa, one could interpret this as a positive association between education and housing affordability, as affordability increases. Outcomes are in line with prior literature (Linneman and Megbolugbe, 1992; DeVaney et al., 2004; Elmelech, 2004, Ben-Shahar et al., 2019). Additionally, outcomes of the consumption-adjusted measure suggest a stronger negative association between education and housing affordability burdens compared to the traditional measure, which is in line with Ben-Shahar et al. (2019). Therefore, Hypothesis II is supported.

The variables, *country* and *ethnicity* show similarities, but are not highly correlated to each other (0,3753) and measure different aspects of household characteristics of the population. For the consumption-adjusted measure, results indicate an 7,9 percent increase in housing affordability burdens among respondents born in a non-western country, and a marginal and insignificant decrease in housing affordability burdens among respondents born in a western country. For the consumption-adjusted measure, results indicate a 5,7 and 0,1 percent increase in housing affordability burdens for first and second generation immigrants respectively. The traditional approach does not recognize significant and marginal coefficients for first and second generation immigrants. Vice versa, the consumption-adjusted measure acknowledges a decrease in housing affordability for respondents born in a non-western country, and first and second generation immigrants. This is in line with prior research (DeVaney et al., 2004; Murdie, 2003; Ben-Shahar et al., 2019). Additionally, outcomes of the consumption-adjusted measure suggest stronger and significant negative associations between *country* and *ethnicity* and housing affordability burdens compared to the traditional method, which is in line with Ben-Shahar et al. (2019). Therefore, Hypothesis III is supported.

Finally, *household composition,* comparable with marital status, shows the association between household composition and housing affordability burdens. The reference category is a household which is single. A decrease of 34 and 47 percent in housing affordability burdens is examined for couples and couples with children compared to a ‘single household’. This is likely due to the accumulation of income. This decrease in housing affordability burdens, or increase in housing affordability, is in line with prior literature (Hendershott et al., 2009; Ben-Shahar et al., 2019). Also, a one parent family experiences less housing affordability burdens compared to the reference group, likely due to a higher average income for

the group of respondents. No underlying reason is found why one parent families experience higher housing affordability compared to single households. At last, a non-family household experiences a 28,4 percent decrease in housing affordability burdens compared to the reference group, this is likely due to households living together under one roof, note that the age of the respondents within this group is generally relatively low, whilst their household income is relatively high. The incomes of this household composition group is most likely combined, and households exists of two adults or more. The differences in the association between household composition and housing affordability burdens are marginal between the consumption-adjusted- and the traditional measure, which is in line with Ben-Shahar et al. (2019). Therefore, Hypothesis IV is accepted.

Overall, the consumption adjusted approach recognizes stronger increases and declines in housing affordability burdens among the dependent variables, whilst simultaneously achieving a higher R-square, essentially indicating a stronger model fit to measure housing affordability burdens within the Netherlands.

4.3 The course of 2009 - 2018 and the relative consumption-adjusted affordability ratio

This chapter presents the relative consumption adjusted measure over the course of the years and its effects several groups of households, of which the main result is focused on the severely affordability burdened groups of the population[7]. Moreover, a chow-test is performed to distinguish differences between urban and rural areas, likewise tested on the severity of the housing affordability of the distinguished groups described in more detail in the discussion of the models.

To measure the severity of a household's affordability burden a new dependent variable is constructed. The estimation is as follows:

$$ln(RHA_{iy}) = \beta_0 + \beta_1 Year + \beta_2 Household\ characteristics_i + \beta_3 LFE_i + \varepsilon_{iy} \qquad \text{eq. (6)}$$

The *i* and *y* denote the indices household and time respectively. The dependent variable RHA presents the relative-to-the-mean measure, being $(HP^{CA} / Income)_{iy} - \underline{(HP^{CA} / Income)_{y}}$, where $\underline{(HP^{CA} / Income)_{y}}$ is the average of $(HP^{CA} / Income)_{iy}$ of all households *i* in year *y* respectively. First the β_0 denotes the intercept of the dependent variable *RHA*, coefficients of the independent variables are indicated by β_1, β_2 and β_3. The independent variables are *Year*, the years 2009, 2012, 2015 and 2015, *household characteristics* stand for the household characteristics used in the main model, being country, ethnicity, age, education and

[7] Note that at first economy wide trends were examined in relation to the RHA measure, being HPI and average income per year. However, as only 4 years are examined, results are rigid and hard to interpret. Although limited, the course of the relative-to-the mean over the years 2009, 2012, 2015 and 2018 is of interest, as it can be translated to the (in)equality of housing affordability in the Netherlands.

households composition. The *LFE* stands for the locational fixed effect, being peripheral indicators and provinces respectively. Additionally, ln(x) stands for the log operator of the variables and ε is the error term in the equation.

Table 6. Outcomes obtained from the regression based on eq. (6).

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Constant	0.794***	-0.244***	1.389***	1.350***
	(0.0422)	(0.0287)	(0.0227)	(0.0111)
2009	Base	Base	Base	Base
2012	-0.0376***	-0.0278***	0.00405	-0.0345***
	(0.0109)	(0.00682)	(0.00591)	(0.00223)
2015	-0.205***	-0.204***	0.121***	-0.342***
	(0.0112)	(0.00701)	(0.00589)	(0.00238)
2018	-0.123***	-0.122***	0.0673***	-0.202***
	(0.0110)	(0.00695)	(0.00598)	(0.00228)
Observations	88,895	129,242	43,627	43,628
Household Characteristics	Included	Included	Included	Included
LFE	Included	Included	Included	Included
Scale	National	National	National	National
R-squared	0.120	0.144	0.096	0.402
Sample	RHA > mean	RHA < mean	Least affordable quantile	Most affordable quantile

Note: dependent variable: the log of the relative-to-the-mean measure (RHA_{iy}). Robust standard errors in parentheses and significance depicted with *** p<0.01, ** p<0.05, * p<0.1 respectively.

Importantly, the *RHA* measures to which extent the household affordability burden exceeds the mean burden of the population. Results are presented in table 6. First the equation is solely estimated for the group that exceed the mean burden, and therefore experience a more than average housing affordability burden, $(HP^{CA} / Income)_{iy} > \underline{(HP^{CA} / Income)_{y}}$, in column 1. Secondly, the group that is least affordability burdened is tested $(HP^{CA} / Income)_{iy} < \underline{(HP^{CA} / Income)_{y}}$ in column 2. As a log operator on negative values is not possible, the *RHA* below the mean is transformed to positives to perform the log operator. Note that the values remain the same, and are simply transformed to positives, therefore no changes in the log occur. Thereafter the sample is stratified in quantiles based on the consumption-adjusted housing affordability

(HP_{CA} / *Income*) of which the first and fifth quantile are of interest, being the highest and lowest housing affordability burdened quantiles respectively.

Column 1 present the association between the years and the relative-to-the-mean, with the year 2009 as the reference category for the severely affordability burdened population of the sample with the addition of controlling for the locational and household characteristics. Results of the estimation indicate a 3,8, 20,5 and 12,3 percent decrease in the relative-to-the-mean consumption adjusted affordability ratio for the most affordability burdened households, for the years 2012, 2015 and 2018 respectively, compared to the reference category. Note that all estimates are significant on a 1 percent level. The decrease in the relative-to-the-mean consumption adjusted affordability ratio is surprising as the expectation are that the lower income groups are part of the most affordability burdened group. On the other hand, following expectations, affordability burdens increase as house prices rise. A likely reason for the decrease in housing affordability burdens when average income rises is due to the average decline of HPI over the years. In other words, the average income increase is likely stronger during these years than the house price indices. Note that once house prices rise rapidly, i.e., stronger in comparison to the average income, the effect is likely to change directions (increase affordability burdens) for the most affordability burdened groups.

Column 2 presents the estimation of equation 6, which is repeated for the least affordability burdened population of the sample $\mathbf{(}HP^{CA} / Income\mathbf{)}_{iy} < \mathbf{(}\underline{HP^{CA} / Income}\mathbf{)}_{\underline{y}}$. For this sub-sample, results of the estimation indicate a 2,8, 20,4 and 12,2 percent decrease in the relative-to-the-mean consumption adjusted affordability ratio for the least affordability burdened households, for the years 2012, 2015 and 2018 respectively, compared to the reference category. This essentially indicates that the least affordability burdened sample benefits more or less the same from an increase in contrast to the most affordability burdened sample and simultaneously experiences less relative housing affordability increases once HPI increases occur. Simply put, the most and least affordability burdened groups both grow more towards the mean during an economic downturn, which is most apparent in 2015.

As described earlier, quantiles related to housing affordability are divided, of which the most- and least affordability burdened groups are examined in column 3 and 4 respectively. Columns 3 shows the estimated results of the most affordability burdened group within the relative consumption-adjusted housing affordability measure. For this sub-sample, results of the estimation indicate a 0,4, 12,2 and 6,7 percent increase in the relative-to-the-mean consumption adjusted affordability ratio for the most affordability burdened quantile, for the years 2012, 2015 and 2018 respectively, compared to the reference category. Note that the direction of the coefficients change for the most affordability burdened quantile, compared to column 1 (most affordability burdened households), indicating an increase further from the mean. This essentially means, during the economic downturn the severely burdened quantile experienced a decrease in housing affordability, compared to mean.

The direction of coefficients in column 4 remain in line with column 2, however the association is stronger. Results indicate a 3,4 34,2 and 20,2 percent decrease in the relative-to-the-mean consumption-adjusted affordability ratio for the least affordability burdened quantile, for the years 2012, 2015 and 2018 respectively, compared to the reference category.

In sum, these results indicate that the overall relative-to-the-mean ratio is decreasing over the years, meaning that there is an elevation in equality between the most and least affordability burdened groups. However, between the top 20% and bottom 20% (column 3 and 4) of the population, an increase(decrease) is noticed for the most(least) affordability burdened groups, for the relative-to-the-mean ratio. Therefore, results are contradictory, indicating that the lowest and highest quantile of the affordability groups are growing apart, whilst the overall housing affordability burdens of the population grow more to one another. Although the contradicting outcome, hypothesis V is rejected, when comparing the total population based on column 1 and 2, with a research aim on the entire population of the Netherlands. On the contrary, Ben-Shahar et al. (2019) found a decline in the quality and consumption adjusted housing affordability in relative terms. They found that the upwards economic trend lead to an elevation of inequality within Israel. Therefore the contractionary results from table 6 might be explained due to the downward economic trend which followed after the GFC between 2009 and 2015, equality elevates, whilst inequality rises after 2015.

4.4 Chow-test, economy wide trends on relative housing affordability, rural vs. urban

This section presents the final models based on equation 6. In table 7, the results are presented. Within this figure, column 1 and 2 show the highest affordability burdened sample, whilst column 3 and 4 illustrates the least affordability burdened sample. A distinction in this model is made between urban and rural areas. The urban periphery is a bundle of households living in (1) an urban center, (2) outside the urban center and (3) green urban. The rural periphery is a bundle of households living in (1) center village and (2) rural areas.

Table 7. Outcomes obtained from the regression based on eq. (6).

VARIABLES	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4
Constant	1.411***	1.372***	1.357***	1.321***
	(0.0369)	(0.0279)	(0.0175)	(0.0138)
2012.Year	-0.0232**	0.0163**	-0.0420***	-0.0315***
	(0.0101)	(0.00733)	(0.00396)	(0.00270)
2015.Year	0.0955***	0.134***	-0.359***	-0.333***
	(0.0101)	(0.00719)	(0.00418)	(0.00291)
2018.Year	0.0169*	0.107***	-0.207***	-0.201***

	(0.00980)	(0.00766)	(0.00386)	(0.00287)
Observations	19,256	24,371	16,930	26,698
Household characteristics	Included	Included	Included	Included
LFE	Included	Included	Included	Included
R-squared	0.078	0.080	0.389	0.411
Periphery	Rural	Urban	Rural	Urban
Sample	Least affordable quantile	Least affordable quantile	Most affordable quantile	Most affordable quantile

Note: dependent variable: the log of the relative-to-the-mean measure (RHA_{iy}). Robust standard errors in parentheses and significance depicted with *** p<0.01, ** p<0.05, * p<0.1 respectively.

Column 1 and 2 distinct the rural and urban samples within the highest affordability burdened quantile. The urban areas within the lowest quantile experience a relatively higher relative-to-the-mean affordability burden compared to the urban areas, indicating that affordability burdens in the urban areas have a stronger association for the severely burdened population compared to the rural areas. Furthermore, column 3 and 4 distinct the rural and urban samples within the lowest affordability burdened quantile. Results are consistent with table 6 where affordability burdens are pronounces within the urban sample, compared to the rural sample. Differences in coefficients between the rural and urban areas for the least affordability burdened groups are marginal. Therefore, no differences are found between rural and urban areas for the least affordability burdened groups.

5. Conclusion

In this study, the consumption-adjusted approach of the study Ben-Shahar et al. (2019)has been examined in comparison to the traditional price-to-income ratio in order to measure potential housing affordability issues in the Netherlands. The traditional measure is criticized for its biases, as it does not account for variability in consumption and preferences of housing among households. The consumption-adjusted measure accounts for these biases and enables to distinguish normative housing affordability within similar household groups. The findings of this thesis could aid policymakers in the Netherlands, as it removes potential individual biases and underlines housing affordability burdens for less privileged households, generally underestimated within the traditional price-to-income measure. which could be translated to policymaking in order to address affordability issues.

The analyses performed, are based on datasets delivered by the ministry of BZK of the Netherlands for the years 2009, 2012, 2015 and 2018. The consumption-adjusted measure revealed more pronounced

affordability burdens among the underprivileged samples, i.e. immigrants, relatively lower educated households, younger adults and single households. Overall inequality between more(less) burdened household groups actually decreases, as both groups grow more towards the average housing affordability burden, likely due to the decrease in housing prices between 2009-2015, as a result of the GFC. This is however not the case for the severely burdened household groups. results indicate stronger economy wide association on the groups that are already heavily affordability burdened, meaning that already severely housing affordability burdened groups experience growing burdens, whilst results for the least affordability burdened groups' burdens decrease. Therefore inequality between the high(low) quantiles of housing affordability burdened groups tend to increase over the years, pressing on the quality and consumption adjusted affordability for lower income households. Policymakers could focus on alleviating affordability problems for the most affordability burdened household groups and the lesser privileged groups. An example could be reducing the transfer tax for these groups to enable them to become homeowners.

There are however some limitations present in this study. First it only accounts for a total of four years with, on average, a gap of three years in between, trends are therefore less subtle, more rigid compared to yearly data. This is especially important within the relative-to-the-mean ratio, as interpretability is hard, reliability is limited due to rigid trends. Second, the substitute for market price, being WOZ-value within this study might deviate from the actual market price at that given time, as WOZ-values are standardized values for types of homes and lag in time. This might result in a distorted housing affordability measure, as it does not represent the 'market price' at that given time. Third, no weighted averages have been used in this study provided by the ministry of BZK.

Although the limitations are specific to this thesis, future research should try to avoid these issues. Additionally, future research could focus on a set of three important considerations. First, a focus on an economic upturn could be interesting to examine with the use of the consumption-adjusted measure, as in this study an overall increase in equality is found during an economic downturn, the effect could potentially be vice versa. This could be interesting for policymaking related to economic cycles. Secondly, the inclusion of mortgage rates could assist in explaining volatility within housing affordability, especially when examining housing affordability within different economic cycles. Thirdly, research could potentially find a way to include general living costs in the consumption-adjusted measure, following the residual income method, and create a mixture of housing affordability measures in order to fully correct for all potential biases, as the consumption-adjusted measure could still potentially inherit biases.

Literature

Ali, A., 2022. Canada needs 5.8 million new homes by 2030 to tackle affordability crisis, CMHC warns. *The Canadian Press.* Available at: https://www.cbc.ca/news/canada/toronto/housing-affordability-cmhc-report-2030-1.6498898 (Accessed 25 June 2022).

Baker, E., Bentley, R., Lester, L. and Beer, A., 2016. Housing affordability and residential mobility as drivers of locational inequality. *Applied Geography*, 72, 65-75.

Beer, A., Baker, E., Wood, G. and Raftery, P., 2011. Housing Policy, Housing Assistance and the Wellbeing Dividend: Developing an Evidence Base for Post-GFC Economies. *Housing Studies*, 26(7), 1171-1192.

Ben-Shahar, D., Gabriel, S. and Golan, R., 2019. Housing affordability and inequality: A consumption-adjusted approach. *Journal of Housing Economics*, 45, 1-12.

Ben-Shahar, D, Gabriel, S. and Golan, R., 2020. Can't get there from here: Affordability distance to a superstar city. *Regional Science and Urban Economics*, 80, 1-15.

Bogdon, A.S. and Can, A., 1997. Indicators of Local Housing Affordability: Comparative and Spatial Approaches. *Journal of Real Estate Economics*, 25(1), 43-80.

Bujang, A. A., Zarin, H. A. and Jumadi, N., 2010. The relationship between demographic factors and housing affordability. *Malaysian Journal of Real Estate*, 5(1), 49-58.

Calcasa, 2020. Calcasa jaarverslag miljoenenwoningen 2020: grootste stijging ooit in aantal miljoenenwoningen. *Calcasa*. Available at: https://www.calcasa.nl/nieuws/calcasa-jaarverslag-miljoenenwoningen-2020-grootste-stijging-ooit-in-aantal-miljoenenwoningen (Accessed 5 January 2022).

CBS, 2020. Grootste huurstijging in zes jaar. *Centraal Bureau voor de Statistiek* Available at: https://www.cbs.nl/nl-nl/nieuws/2020/37/grootste-huurstijging-in-zes-jaar (Accessed 10 December 2021).

CBS, 2021. Voorraad woningen; eigendom, type verhuurder, bewoning, regio. Available at: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82900NED/table?dl=10740 (Accessed 8 August 2022).

Clark, W.A.V., Deurloo, M.C. and Dieleman, F.M., 2000. Housing Consumption and Residential Crowding in U.S. Housing Markets. *Journal of Urban Affairs*, 22(1), 49-63.

DeVaney, S.A., Chiremba, S. and Vincent, A., 2004. Life Cycle Stage and Housing Cost Burden. *Financial Counseling and Planning*, 15(1), 31-39.

Dewita, Y., Yen, B.T.H. and Burke, M., 2018. The effect of transport cost on housing affordability: Experiences from the Bandung Metropolitan Area, Indonesia. *Land Use Policy*, 79, 507-519.

Elemech, Y., 2004. Housing Inequality in New York City: Racial and Ethnic Disparities in Homeownership and Shelter-Cost Burden. *Housing, Theory and Society*, 21(4), 163-175.

Europees Parlement, 2021. Europees Parlement pleit voor maatregel om huisvestingscrisis op te lossen. *Europees Parlement Nieuws*. Available at: https://www.europarl.europa.eu/news/nl/headlines/society/20210114STO95643/europees-parlement-pleit-voor-maatregelen-om-huisvestingscrisis-op-te-lossen (Accessed 10 December 2021).

Eurostat, 2022. *Share of housing costs in disposable household income, by type of household and income group – EU-SILC survey*. Eurostat. Available at: https://ec.europa.eu/eurostat/databrowser/view/ILC_MDED01__custom_1514507/bookmark/table?lang=en&bookmarkId=50a09682-7e88-4f27-9e19-125916e1cace (Accessed 25 March 2022).

Fisher, L.M., Pollakowski, H.O. and Zabel, J., 2009. Amenity-based Housing Affordability Indexes. *Real Estate Economics*, 37(4), 705-746.

Gan, Q. and Hill, R.J., 2009. Measuring housing affordability: Looking beyond the median. *Journal of Housing Economics*, 18(2), 115-125.

Gyourko, J. and Linneman, P., 1993. Thee Affordability of the American Dream: An Examination of the Last 30 Years. *Journal of Housing Research*, 4(1), 39-72.

Haffner, M.E.A. and Boumeester, H.J.F.M., 2010. The Affordability of Housing in the Netherlands: An Increasing Income Gap Between Renting and Owning? *Housing Studies*, 25(6), 799-820.

Haffner, M.E.A. and Boumeester, H.J.F.M., 2014. Is renting unaffordable in the Netherlands? *International Journal of Housing Policy*, 14(2), 117-140.

Haffner, M.E.A. and Boumeester, H.J.F.M., 2015. Housing affordability in the Netherlands: the impact of rent and energy costs. *Journal of Housing and the Built Environment*, 30(2), 293-312.

Haffner, M.E.A. and Hulse, K., 2021. A fresh look at contemporary perspectives on urban housing affordability. *International Journal of Urban Sciences*, 25(1), 59-79.

Haffner, M.E.A. and Heylen, K., 2011. User Costs and Housing Expenses. Towards a more Comprehensive Approach to Affordability. *Housing Studies*, 26(4), 593-614.

Hancock, K., 1993. Can pay – won't pay – or economic-principles of affordability. *Urban Studies*, 30(1), 127-145.

Hendershott, P.H., Ong, R., Wood, G.A. and Glatau, P., 2009. Marital history and home ownership: Evidence from Australia. *Journal of Housing Economics*, 18(1), 13-24.

Heylen, K., 2014. Housing affordability and the effect of housing subsidies. *Katholieke Universiteit Leuven faculteit Sociale Wetenschappen*, Proefschrift.

Hoek van, T., Koning, M., Bruinsma, H. and Uunk, V., 2020. Betaalbaarheid op de woningmarkt. *Economisch Instituut voor de Bouw*, 1-52.

Hulchanski, J.D., 1995. The concept of housing affordability: Six contemporary uses of the housing expenditure-to-income ratio. *Housing Studies*, 10(4), 471-491.

Herbaugh, A., 2022. Housing affordability continues to plunge in California. *KSBY*. Available at: https://www.ksby.com/news/price-of-paradise/housing-affordability-continues-to-plunge-in-california (Accessed 25 June 2022).

Kim, K. and Cho, M., 2010. Structural Changes, Housing Price Dynamics and Housing Affordability in Korea. *Housing Studies*, 25(6), 839-856.

Kutty, N.K., 2005. A new measure of housing affordability: Estimates and analytical results. *Housing Policy Debate*, 16(1), 113-142.

Lerman, D.L. and Reeder, W.J., 1987. The Affordability of Adequate Housing. *Real Estate Economics*, 15(4), 389-404.

Li, J., 2014. *Recent trends on housing affordability research: where are we up to?* Working Paper No. 5. https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID2555439_code2148731.pdf?abstractid=2555439&mirid=1

Lin, Y., Chang, C. and Chen, C., 2014 Why homebuyers have a high housing affordability problem: Quantile regression analysis. *Habitat International*, 43, 41-47.

Linneman, P.D. and Megbolugbe, I.F., 1992. Housing Affordability: Myth of Reality? *Urban Studies*, 3(4), 369-392.

Mayer, C.J. and Engelhardt, G.V., 1996. Gifts, Down Payments, and Housing Affordability. *Journal of Housing Research*, 7(1), 59-77.

Murdie, R.A., 2003. Housing affordability and Toronto's rental market: perspectives from the housing careers of Jamaican, Polish and Somali Newcomers. *Housing, Theory and Society*, 20(4), 183-196.

Newman, S.J. and Holupka, C.S., 2014. Housing Affordability and Child Well-Being. *Housing Policy Debate*, 25(1), 116-151.

Norris, M. And Shiels, P., 2007. Housing Affordability in the Republic of Ireland: Is Planning Part of the Problem or Part of the Solution? *Housing Studies*, 22(1), 45-61.

NOS, 2021. Ruim half miljoen huishoudens leven in 'energiearmoede'. *Nederlandse Omroep Stichting.* Available at: https://nos.nl/artikel/2398887-ruim-half-miljoen-huishoudens-leven-in-energiearmoede (Accessed 12 November 2021).

Quigley, J.M. and Raphael, S., 2004. Is housing unaffordable? Why isn't it more affordable? *Journal of Economic Perspectives*, 18(1), 191-214.

Rijksoverheid, 2021. Welke regels gelden er voor een huurverhoging? *Rijksoverheid*. Available at: https://www.rijksoverheid.nl/onderwerpen/woning-huren/vraag-en-antwoord/welke-regels-gelden-er-voor-een-huurverhoging (Accessed 10 December 2021).

Rowley, S. and Ong, R., 2012. *Housing affordability, housing stress and household wellbeing in Australia*. Melbourne: Australian Housing and Urban Research Institute.

Sawhill, I., 2018. *The forgotten Americans: An economic agenda for a divided nation*. New Haven: Yale University Press.

Skaburskik, A. 1997. Gender Differences in Housing Demand. *Urban Studies*, 34(2), 275-320.

Skaburskis, A., 2004. Decomposing Canada's Growing Housing Affordability Problem: Do City Differences Matter? *Urban Studies*, 41(1), 117-149.

Stone, M.E., 2006. What is housing affordability? The case for the residual income approach. *Housing Policy Debate*, 17(1), 151-184.

Thalmann, P., 1999. Identifying households which need housing assistance. *Urban Studies*, 36(11), 1933-1947.

Turner, T.M. and Luea, H., 2009. Homeownership, wealth accumulation and income status. *Journal of Housing Economics*, 18, 104-114.

Wetzstein, S., 2017. The global urban housing affordability crisis. *Urban Studies*, 54(14), 3159-3177.

Wetzstein, S., 2021. Toward Affordable Cities? Critically Exploring the Market-Based Housing Supply Policy Proposition. *Housing Policy Debate*, 32(3), 506-532.

Whitehead, C.M.E., 1991. From need to affordability – an analysis of UK housing objectives. *Urban Studies*, 28(6), 871-887.

Wilcox, S., 2006. *The geography of affordable and unaffordable housing*. [online]. York: York Publishing Services Ltd. Available at: https://www.jrf.org.uk/file/37221/download?token=KusDCLoY&filetype=full-report (Accessed 20 June 2022).

Zhang, C., Jia, S. and Yang, R., 2016. Housing affordability and housing vacancy in China: The role of income inequality. *Journal of Housing Economics*, 33, 4-14.

Zyed, Z.A.S., Aziz, W.N.A.W.A.A. and Hanif, N.R., 2016. Housing Affordability Problems among Young Households. *Journal of Surveying, Constuction and Property (JSCP)*, 7(1), 1-18.

Appendices

Appendix A, computation of estimated coefficients of equation (2)

VARIABLES	(1) Groningen	(2) Friesland	(3) Drenthe	(4) Overijssel	(5) Flevoland	(6) Gelderland
Constant	11.6218***	11.6323***	11.6263***	11.6766***	11.8696***	11.8306***
	(0.0182)	(0.0184)	(0.0232)	(0.0103)	(0.0161)	(0.0073)
Surface	0.0038***	0.0039***	0.0040***	0.0041***	0.0042***	0.0036***
	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Building_Age	-0.0008***	-0.0017***	0.0002	-0.0016***	-0.0088***	0.0002***
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0003)	(0.0001)
Stories	0.0544***	0.0330***	0.0121	0.0309***	-0.0104**	0.0394***
	(0.0061)	(0.0051)	(0.0077)	(0.0029)	(0.0041)	(0.0022)
DumNew	0.1339**	0.0628	0.1469***	0.0642***	-0.0698**	0.0262
	(0.0536)	(0.0478)	(0.0525)	(0.0214)	(0.0292)	(0.0162)
Observations	5,140	5,018	3,927	14,198	5,941	32,811
R-squared	0.3879	0.4038	0.3881	0.3791	0.4722	0.3445
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

VARIABLES	(7) Utrecht	(8) Noord-Holland	(9) Zuid- Holland	(10) Zeeland	(11) Noord-Brabant	(12) Limburg
Constant	11.6501***	11.7336***	11.5049***	11.7440***	11.8539***	11.6302***
	(0.0100)	(0.0069)	(0.0044)	(0.0132)	(0.0082)	(0.0113)
Surface	0.0047***	0.0049***	0.0058***	0.0040***	0.0041***	0.0041***
	(0.0001)	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)
Building_Age	0.0013***	0.0016***	-0.0015***	-0.0016***	-0.0012***	-0.0020***
	(0.0001)	(0.0001)	(0.0000)	(0.0001)	(0.0001)	(0.0001)
Stories	0.0835***	0.0272***	0.0672***	-0.0016	0.0252***	0.0361***
	(0.0031)	(0.0017)	(0.0012)	(0.0040)	(0.0024)	(0.0030)
DumNew	0.0845***	0.1192***	0.0514***	-0.0296	0.0349**	0.0823**
	(0.0205)	(0.0197)	(0.0113)	(0.0340)	(0.0175)	(0.0373)
Observations	12,614	26,244	64,190	10,564	25,673	11,817
R-squared	0.4773	0.3777	0.5048	0.3199	0.4238	0.4191
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix B

Matrix of correlations

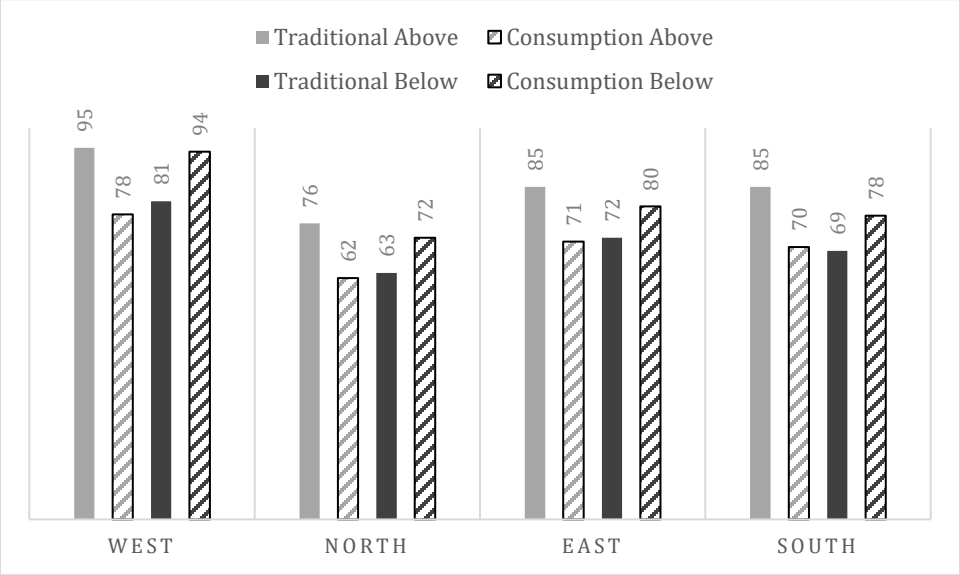
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) lnHA_CA	1.000								
(2) Country_Born	0.025	1.000							
(3) Ethnicity_Gene~n	0.012	0.376	1.000						
(4) AGE	0.245	-0.049	-0.089	1.000					
(5) Education	-0.343	0.015	0.020	-0.294	1.000				
(6) Partner	-0.410	-0.025	-0.036	-0.114	0.068	1.000			
(7) Household_Comp~n	-0.326	0.063	0.064	-0.351	0.087	0.372	1.000		
(8) Periphery	0.036	-0.119	-0.132	0.080	-0.054	0.149	0.052	1.000	
(9) Year	-0.164	-0.031	-0.018	0.058	0.070	-0.009	-0.013	0.107	1.000

Appendix C

Variance inflation factor

	VIF	1/VIF
Household Composition	1.318	.759
Age	1.263	.792
Partner	1.191	.84
Ethnicity Generation	1.184	.844
Country Born	1.175	.851
Education	1.11	.901
Periphery	1.07	.934
Year	1.025	.975
Province	1.006	.994
Mean VIF	1.15	

Appendix D



Traditional- and consumption-adjusted approach on regional level within the Netherlands. Housing affordability measure in Months. N = 51,807

Appendix E: stata commands

```
*DATASET 2009 - DATAPREP
import spss using "C:\Users\shawn\Documents\Housing Affordability\Data\Raw
data\WoON2009_e_1.5.sav"

*Generate new var to indicate year
generate Year_2009 = 2009
label variable Year_2009 "2009"

*Data prep voor dataset 2009
keep aantkind gemcode hwmbrt stedgem Cohesie vromhh tonderho ruimte opptbin opptbin2 aantalpp
BhVorm voplop vltoplpa etniop3 etnipa3 gblop3 etnigop gslop Leeftijd partner samhh5 aankprs bjaar
jrgkoch2 jrkomwon Woonvrd Bestwon vorm HYP BKOOP prov waarwon Year_2009

*Rename variables
ren aantkind Children
ren gemcode Municipality
ren hwmbrt Periphery
ren vromhh Income
ren opptbin Surface
ren aantalpp People
ren BhVorm Tenant
ren voplop Education
ren etniop3 Ethnicity
ren gblop3 Country_Born
ren etnigop Ethnicity_Generation
ren gslop Gender
ren Leeftijd AGE
ren partner Partner
ren samhh5 Household_Composition
ren bjaar Building_Age
ren jrgkoch2 Year_Bought
ren jrkomwon Year_Since_Tenancy
ren Woonvrd Stories
ren Bestwon Newly_Built
ren vorm Type_Home
ren prov Province
ren waarwon WOZ
ren Year_2009 Year

*recode Education --> https://www.cbs.nl/nl-nl/nieuws/2019/33/verschil-levensverwachting-hoog-en-
laagopgeleid-groeit/opleidingsniveau
sum Education
drop if Education==11
```

```
recode Education (1/5=1) (6/8=2) (9/10=3)
label variable Education "Education level"
label define Education_label 1 "Low educated" 2 "Secondary educated" 3 "Highly educated"
label values Education Education_label

*recode Periphery
sum Periphery
label variable Periphery "Periphery level"
label define Periphery_label 1 "Urban center" 2 "Urban outside center" 3 "Green Urban" 4 "Center
Village" 5 "Rural"
label values Periphery Periphery_label

*recode Ethnicity
sum Ethnicity
label variable Ethnicity "Ethnicity level"
label define Ethnicity_label 1 "Dutch" 2 "non Western" 3 "Western"
label values Ethnicity Ethincity_label

*recode Type_Home
sum Type_Home
label variable Type_Home "Type_Home level"
label define Type_Home_label 1 "single-family home" 2 "multi-family home"
label values Type_Home Type_Home_label

*recode Gender
sum Gender
label variable Gender
label variable Gender "Gender type"
label define Gender_label 1 "Man" 2 "Female"
label values Gender Gender_label

*recode Children
sum Children
recode Children (0=0) (1=1) (2=2) (3/8=3)
label variable Children "Number of Children"
label define Children_label 0 "No Children" 1 "1 Child" 2 "2 Children" 3 "3 Children or more"
label values Children Children_label

*Create Age - age of structure at the time of tenancy
replace Year_Since_Tenancy = 2009-Building_Age
label variable Year_Since_Tenancy "Age Building"

*Generate number of adults
gen Adults = People-Children
```

```
*Gen Newly Built
gen DumNew = Year-Building_Age
recode DumNew (0/2=1) (3/max=0)

*drop / what to drop
drop if Adults > 5
keep if !missing(WOZ)
keep if !missing(Surface)
keep if !missing(Building_Age)
keep if !missing(Type_Home)
keep if !missing(Ethnicity_Generation)
*based on Staatscourant, 2009, drop incomes under € 10.878,-
drop if Income < 10878

*tab cluster Adults / Children
tab Adults Children
tab Adults Children, cell nofreq

*WOZ waardes - *1000
replace WOZ = WOZ*1000
winsor2 WOZ, replace cuts(1 99)

save file

*DATASET 2012 - DATAPREP
import spss using "C:\Users\shawn\Dropbox\My PC (DESKTOP-EPK7QET)\Documents\Real
Estate\Thesis\Housing Affordability\Data\Datasets\Raw data\WoON2012_e_1.1.sav"

*Generate new var to indicate year
generate Year_2012 = 2012
label variable Year_2012 "2012"

*Data prep voor dataset 2012
keep AANTKIND GemCode hwmbrt stedgem Cohesie vromhh Tonderho Kamers ruimte OppTBin
OppTBin2 AantalPP huko VOplOP VOplPA etniop3 etnipa3 gblop3 etnigop gslop Leeftijd partner
SAMHH5 AankPrs BJaar JrGKoch2 JrKomWon WoonVrd BestWon vorm hyp bkoop khuurii nhuur
g4_2 g4_3 prov wozwaarde Year_2012

*Rename variables
ren AANTKIND Children
ren GemCode Municipality
ren hwmbrt Periphery
ren vromhh Income
ren Kamers Rooms
ren OppTBin Surface
ren AantalPP People
ren huko Tenant
```

```
ren VOplOP Education
ren etniop3 Ethnicity
ren gblop3 Country_Born
ren etnigop Ethnicity_Generation
ren gslop Gender
ren Leeftijd AGE
ren partner Partner
ren SAMHH5 Household_Composition
ren BJaar Building_Age
ren JrGKoch2 Year_Bought
ren JrKomWon Year_Since_Tenancy
ren WoonVrd Stories
ren BestWon Newly_Built
ren vorm Type_Home
ren prov Province
ren wozwaarde WOZ
ren Year_2012 Year

*recode Education --> https://www.cbs.nl/nl-nl/nieuws/2019/33/verschil-levensverwachting-hoog-en-laagopgeleid-groeit/opleidingsniveau
sum Education
drop if Education==11
recode Education (1/5=1) (6/8=2) (9/10=3)
label variable Education "Education level"
label define Education_label 1 "Low educated" 2 "Secondary educated" 3 "Highly educated"
label values Education Education_label

*recode Periphery
sum Periphery
label variable Periphery "Periphery level"
label define Periphery_label 1 "Urban center" 2 "Urban outside center" 3 "Green Urban" 4 "Center Village" 5 "Rural"
label values Periphery Periphery_label

*recode Periphery_Mun
sum Periphery_Mun
label variable Periphery_Mun "Periphery_Mun level"
label define Periphery_Mun_label 1 "Very Urban" 2 "Urban" 3 "Moderately Urban" 4 "Little Urban" 5 "Not Urban"
label values Periphery_Mun Periphery_Mun_label

*recode Ethnicity
sum Ethnicity
label variable Ethnicity "Ethnicity level"
label define Ethnicity_label 1 "Dutch" 2 "non Western" 3 "Western"
label values Ethnicity Ethincity_label

*recode Type_Home
sum Type_Home
label variable Type_Home
label variable Type_Home "Type_Home level"
```

```
label define Type_Home_label 1 "single-family home" 2 "multi-family home"
label values Type_Home Type_Home_label

*recode Gender
sum Gender
label variable Gender
label variable Gender "Gender type"
label define Gender_label 1 "Man" 2 "Female"
label values Gender Gender_label

*recode Children
sum Children
recode Children (0=0) (1=1) (2=2) (3/6=3)
label variable Children "Number of Children"
label define Children_label 0 "No Children" 1 "1 Child" 2 "2 Children" 3 "3 Children or more"
label values Children Children_label

*Create Age - age of structure at the time of tenancy
replace Year_Since_Tenancy = 2012-Building_Age
label variable Year_Since_Tenancy "Age Building"

*Generate number of adults
gen Adults = People-Children

*drop / what to drop
drop if Adults > 5
keep if !missing(WOZ)
keep if !missing(Rooms)
keep if !missing(Building_Age)
keep if !missing(Type_Home)
keep if !missing(Ethnicity_Generation)
*drop households if income is below 11.229, based on Staatscourant, 2012
drop if Income < 11229

*create dummy for Age Building
gen DumNew = Year-Building_Age
recode DumNew (0/2=1) (3/max=0)

*WOZ waardes - *1000
winsor2 WOZ, replace cuts(1 99)

save file

*DATASET 2015 - DATAPREP
use "C:\Users\shawn\Dropbox\My PC (DESKTOP-EPK7QET)\Documents\Real Estate\Thesis\Housing
Affordability\Data\Raw data\WoON2015_e_1.0.dta"

*Generate new var to indicate year
generate Year_2015 = 2015
label variable Year_2015 "2015"
```

```
*Data prep voor dataset 2015 ** SRTWONING / SRTAPP / HUISTYP MIST (!)
keep AantKind GemCode hwmbrt stedgem Cohesie vromhh Tonderho Kamers Ruimte gebruiksopp
OppWon7 AantalPP huko NivBehOP NivBehPA etniop3 etnipa3 gblop3 etnigop gslop Leeftijd Partner
SamHH5 AankPrs bjaarbag JrGekocht JrKomWon WoonVrd BestndWon vorm hyp bkoop khuurii nhuur
g4_2 g4_3 prov WOZwaarde Year_2015

*Rename variables
ren AantKind Children
ren GemCode Municipality
ren hwmbrt Periphery
ren vromhh Income
ren gebruiksopp Surface
ren AantalPP People
ren huko Tenant
ren NivBehOP Education
ren etniop3 Ethnicity
ren gblop3 Country_Born
ren etnigop Ethnicity_Generation
ren gslop Gender
ren Leeftijd AGE
ren Partner Partner
ren SamHH5 Household_Composition
ren bjaarbag Building_Age
ren JrGekocht Year_Bought
ren JrKomWon Year_Since_Tenancy
ren WoonVrd Stories
ren BestndWon Newly_Built
ren vorm Type_Home
ren prov Province
ren WOZwaarde WOZ
ren Year_2015 Year

*recode Education --> https://www.cbs.nl/nl-nl/nieuws/2019/33/verschil-levensverwachting-hoog-en-
laagopgeleid-groeit/opleidingsniveau
sum Education
drop if Education==11
recode Education (1/5=1) (6/8=2) (9/10=3)
label variable Education "Education level"
label define Education_label 1 "Low educated" 2 "Secondary educated" 3 "Highly educated"
label values Education Education_label

*recode Periphery
sum Periphery
label variable Periphery "Periphery level"
label define Periphery_label 1 "Urban center" 2 "Urban outside center" 3 "Green Urban" 4 "Center
Village" 5 "Rural"
label values Periphery Periphery_label

*recode Periphery_Mun
sum Periphery_Mun
label variable Periphery_Mun "Periphery_Mun level"
```

```
label define Periphery_Mun_label 1 "Very Urban" 2 "Urban" 3 "Moderately Urban" 4 "Little Urban" 5 "Not Urban"
label values Periphery_Mun Periphery_Mun_label

*recode Ethnicity
sum Ethnicity
label variable Ethnicity "Ethnicity level"
label define Ethnicity_label 1 "Dutch" 2 "non Western" 3 "Western"
label values Ethnicity Ethincity_label

*recode Type_Home
sum Type_Home
label variable Type_Home "Type_Home level"
label define Type_Home_label 1 "single-family home" 2 "multi-family home"
label values Type_Home Type_Home_label

*recode Gender
sum Gender
label variable Gender
label variable Gender "Gender type"
label define Gender_label 1 "Man" 2 "Female"
label values Gender Gender_label

*recode Children
sum Children
recode Children (0=0) (1=1) (2=2) (3/7=3)
label variable Children "Number of Children"
label define Children_label 0 "No Children" 1 "1 Child" 2 "2 Children" 3 "3 Children or more"
label values Children Children_label

*Create Age - age of structure at the time of tenancy
replace Year_Since_Tenancy = 2015-Building_Age
label variable Year_Since_Tenancy "Age Building"

*Generate number of adults
gen Adults = People-Children

*tab clusters
tab Adults
tab Children
tab Periphery

*Gen Newly Built
gen DumNew = Year-Building_Age
recode DumNew (0/2=1) (3/max=0)

*drop / what to drop
drop if Adults > 5
keep if !missing(WOZ)
keep if !missing(Surface)
keep if !missing(Building_Age)
```

```
keep if !missing(Type_Home)
keep if !missing(Ethnicity_Generation)
*drop households if income is below 11.551, based on Staatscourant, 2015
drop if Income < 11551

*WOZ waardes - *1000
winsor2 WOZ, replace cuts(1 99)

save file

*DATASET 2018 - DATAPREP
use "C:\Users\shawn\Dropbox\My PC (DESKTOP-EPK7QET)\Documents\Real Estate\Thesis\Housing Affordability\Data\Datasets\Raw data\WoON2018_e_1.0.dta"

*Generate new var to indicate year
generate Year_2018 = 2018
label variable Year_2018 "2018"

*Data prep voor dataset 2018 *AANTPP IN 5 KLASSEN *GEEN GEM, MAAR STRATUM *GEEN GENDER **WOONOPPERVLAKTE IN 7  *WOONTYPE 6 KLASSEN * CHECK HUKO, LIJKT OMGEKEERD
keep aantkind4 hwmbrt stedgem stratum cohesie vromhh_r tonderho kamers ruimte gebruiksopp oppwon7 aantalpp5 huko nivbehop1 nivbehpa01 etniop3 etnipa3 gblop3 etnigop leeftijd partner samhh5 aankprs bjaarbag jrgekocht jrkomwon woonvrd bestndwon vorm hyp_r bkoopw_r khuurii_r nhuur_r g4_2 g4_3 prov wozwaarde Year_2018 nivbehop2 nivbehop3 nivbehop4 nivbehop5 nivbehop6 nivbehop7 nivbehop8 nivbehop9 nivbehop10 nivbehop11 nivbehop12 nivbehop13 nivbehop14 nivbehop15 nivbehop16 nivbehop17 nivbehpa02 nivbehpa03 nivbehpa04 nivbehpa05 nivbehpa06 nivbehpa07 nivbehpa08 nivbehpa09 nivbehpa10 nivbehpa11 nivbehpa12 nivbehpa13 nivbehpa14 nivbehpa15 nivbehpa16 nivbehpa17

ren aantkind4 Children
ren hwmbrt Periphery
ren stedgem Periphery_Mun
ren stratum Municipality
ren cohesie Cohesion_Neighborhood
ren vromhh_r Income
ren gebruiksopp Surface
ren aantalpp5 People
ren huko Tenant
ren nivbehop1 Education_DUMMY1
ren etniop3 Ethnicity
ren gblop3 Country_Born
ren etnigop Ethnicity_Generation
ren leeftijd AGE
ren partner Partner
ren samhh5 Household_Composition
ren aankprs Transaction_Price
ren bjaarbag Building_Age
ren jrgekocht Year_Bought
ren jrkomwon Year_Since_Tenancy
ren woonvrd Stories
```

```
ren bestndwon Newly_Built
ren vorm Type_Home
ren prov Province
ren wozindex17 WOZ
ren Year_2018 Year
ren nivbehop2 Education_DUMMY2
ren nivbehop3 Education_DUMMY3
ren nivbehop4 Education_DUMMY4
ren nivbehop5 Education_DUMMY5
ren nivbehop6 Education_DUMMY6
ren nivbehop7 Education_DUMMY7
ren nivbehop8 Education_DUMMY8
ren nivbehop9 Education_DUMMY9
ren nivbehop10 Education_DUMMY10
ren nivbehop11 Education_DUMMY11
ren nivbehop12 Education_DUMMY12
ren nivbehop13 Education_DUMMY13
ren nivbehop14 Education_DUMMY14
ren nivbehop15 Education_DUMMY15
ren nivbehop16 Education_DUMMY16
ren nivbehop17 Education_DUMMY17

*recode Education, transform dummies into categories https://www.cbs.nl/nl-nl/nieuws/2019/33/verschil-levensverwachting-hoog-en-laagopgeleid-groeit/opleidingsniveau
gen Education=0
replace Education=1 if Education_DUMMY1==1
replace Education=2 if Education_DUMMY2==1
replace Education=3 if Education_DUMMY3==1
replace Education=4 if Education_DUMMY4==1
replace Education=5 if Education_DUMMY5==1
replace Education=6 if Education_DUMMY6==1
replace Education=7 if Education_DUMMY7==1
replace Education=8 if Education_DUMMY8==1
replace Education=9 if Education_DUMMY9==1
replace Education=10 if Education_DUMMY10==1
replace Education=11 if Education_DUMMY11==1
replace Education=12 if Education_DUMMY12==1
replace Education=13 if Education_DUMMY13==1
replace Education=14 if Education_DUMMY14==1
replace Education=15 if Education_DUMMY15==1
replace Education=16 if Education_DUMMY16==1
drop if Education_DUMMY17==1
recode Education (1/4=1) (5/10=2) (11/16=3)
label variable Education "Education level"
label define Education_label 1 "Low educated" 2 "Secondary educated" 3 "Highly educated"
label values Education Education_label

*recode Periphery
sum Periphery
label variable Periphery "Periphery level"
label define Periphery_label 1 "Urban center" 2 "Urban outside center" 3 "Green Urban" 4 "Center
```

```
Village" 5 "Rural"
label values Periphery Periphery_label

*recode Ethnicity
sum Ethnicity
label variable Ethnicity "Ethnicity level"
label define Ethnicity_label 1 "Dutch" 2 "non Western" 3 "Western"
label values Ethnicity Ethincity_label

*recode Type_Home
sum Type_Home
label variable Type_Home "Type_Home level"
label define Type_Home_label 1 "single-family home" 2 "multi-family home"
label values Type_Home Type_Home_label

*recode Children
sum Children
recode Children (0=0) (1=1) (2=2) (3=3)
label variable Children "Number of Children"
label define Children_label 0 "No Children" 1 "1 Child" 2 "2 Children" 3 "3 Children or more"
label values Children Children_label

*Create Age - age of structure at the time of tenancy
replace Year_Since_Tenancy = 2018-Building_Age
label variable Year_Since_Tenancy "Age Building"

*Generate number of adults
gen Adults = People-Children

*Gen Newly Built
gen DumNew = Year-Building_Age
recode DumNew (0/2=1) (3/max=0)

*drop / what to drop
drop if Adults > 5
keep if !missing(WOZ)
keep if !missing(Rooms)
keep if !missing(Building_Age)
keep if !missing(Type_Home)
keep if !missing(Ethnicity_Generation)
*drop households if income is below 11.958, based on Rijksoverheid, 2018
drop if Income < 11958

*WOZ waardes - *1000
*gen WOZ_1000 = WOZ*1000
winsor2 WOZ, replace cuts(1 99)

save file
```

```
*Affordability Thesis
clear all

*Installs
*ssc install regcheck
*ssc install winsor2
*ssc install r2_a

*Directory
*cd "C:\Users\shawn\Dropbox\My PC (DESKTOP-EPK7QET)\Documents\Real Estate\Thesis\Housing Affordability"
*laptop
*cd "C:\Users\shawn\Documents\Scriptie\Housing Affordability\Data"

set more off
log using Thesis_Record.log, replace text
*computer
import spss using "C:\Users\shawn\Dropbox\My PC (DESKTOP-EPK7QET)\Documents\Real Estate\Thesis\Housing Affordability\Data\WoON2009_e_1.5.sav"
import spss using "C:\Users\shawn\Dropbox\My PC (DESKTOP-EPK7QET)\Documents\Real Estate\Thesis\Housing Affordability\Data\WoON2012_e_1.1.sav"

*MERGE DATASETS
use "C:\Users\shawn\Dropbox\My PC (DESKTOP-EPK7QET)\Documents\Real Estate\Thesis\Housing Affordability\Data\Final datasets\AWoON2009 final.dta"
append using "C:\Users\shawn\Dropbox\My PC (DESKTOP-EPK7QET)\Documents\Real Estate\Thesis\Housing Affordability\Data\Final datasets\AWoON2012 final.dta" "C:\Users\shawn\Dropbox\My PC (DESKTOP-EPK7QET)\Documents\Real Estate\Thesis\Housing Affordability\Data\Final datasets\AWoON2015 final.dta" "C:\Users\shawn\Dropbox\My PC (DESKTOP-EPK7QET)\Documents\Real Estate\Thesis\Housing Affordability\Data\Final datasets\AWoON2018 final.dta"

*drop all Education Dummies
drop Education_DUMMY1
drop Education_DUMMY2
drop Education_DUMMY3
drop Education_DUMMY4
drop Education_DUMMY5
drop Education_DUMMY6
drop Education_DUMMY7
drop Education_DUMMY8
drop Education_DUMMY9
drop Education_DUMMY10
drop Education_DUMMY11
drop Education_DUMMY12
drop Education_DUMMY13
drop Education_DUMMY14
drop Education_DUMMY15
drop Education_DUMMY16
drop Education_DUMMY17
```

```
drop Education_Partner_DUMMY1
drop Education_Partner_DUMMY2
drop Education_Partner_DUMMY3
drop Education_Partner_DUMMY4
drop Education_Partner_DUMMY5
drop Education_Partner_DUMMY6
drop Education_Partner_DUMMY7
drop Education_Partner_DUMMY8
drop Education_Partner_DUMMY9
drop Education_Partner_DUMMY10
drop Education_Partner_DUMMY11
drop Education_Partner_DUMMY12
drop Education_Partner_DUMMY13
drop Education_Partner_DUMMY14
drop Education_Partner_DUMMY15
drop Education_Partner_DUMMY16
drop Education_Partner_DUMMY17

save file

*start with merged dataset
use "C:\Users\shawn\Dropbox\My PC (DESKTOP-EPK7QET)\Documents\Real Estate\Thesis\Housing Affordability\Data\Final datasets\Merged Datasets.dta"

*gen Log WOZ / winsor surface
gen lnWOZ = ln(WOZ)
winsor2 Surface, replace cuts(1 99)

*Building_Age
gen Building_Age2 = Year - Building_Age
sum Building_Age2
drop Building_Age
rename Building_Age2 Building_Age
winsor2 Building_Age, replace cuts(0 99)

*ACLY (Adults Children Periphery Year) stratification - and removal of all strata's with less than 20 obs
egen strata=group(Adults Children Periphery Year)
bysort strata: gen strata_freq = _N
bysort strata: egen strata_min = min(strata_freq)
drop if strata_min <= 20

*equation (1) surface - hoogste R2 (0,3901)
bysort strata: egen totalsurface2=total(Surface)
bysort strata: gen NS_CA_2 = totalsurface2 / strata_freq

*Model with Surface
//Housing affordability - Surface
use "C:\Users\shawn\Documents\Housing Affordability\Data\FINAL\Actual\Datasets\Per Province (eq.2)\Eq1-5 done (actual measure)\Pre-Split.dta"

*Observations per Province
```

```
bysort Province: gen province_freq = _N

keep if Province==1/12

drop _Nobs _R2 _adjR2 _b_Building_Age _b_cons _b_DumNew _b_Stories _b_Surface totalage
AvgBuildingAge totalstories AvgStories totalnew AvgNew

*equation (2) trials - with surface or surface_categorized (higher R2) / asreg
bys Year: asreg lnWOZ Surface Building_Age Stories DumNew
reg lnWOZ Surface Building_Age Stories DumNew i.Year
outreg2 using Provinces.doc, append ctitle(Provincie) dec (4) keep (Surface Building_Age Stories
DumNew) addtext (Year FE, Yes)
regress, coeflegend noheader
gen bcons = _b[_cons]
gen bsurf = _b[Surface]
gen bage = _b[Building_Age]
gen bstory = _b[Stories]
gen bnew = _b[DumNew]
gen b2009 = 0
gen b2012 = _b[2012.Year]
gen b2015 = _b[2015.Year]
gen b2018 = _b[2018.Year]
predict e, res
rename e errorterm

replace b2009 = b2012 if Year==2012
replace b2009 = b2015 if Year==2015
replace b2009 = b2018 if Year==2018
rename b2009 bYear

*equation (3) preparation average characteristics
egen totalage = total(Building_Age)
gen AvgBuildingAge = totalage / province_freq

egen totalstories = total(Stories)
gen AvgStories = totalstories / province_freq

egen totalnew = total(DumNew)
gen AvgNew = totalnew / province_freq

*step (4) equation (3) trials - hedonic pricing model Trials
bysort strata: gen CA_WOZ = exp(bcons + bsurf*(NS_CA_2) + bage*(AvgBuildingAge) +
bstory*(AvgStories) + bYear + errorterm/2)

*step (5) - equation 5 first
bysort strata: gen TR_WOZ = exp(bcons + bsurf*(Surface) + bage*(AvgBuildingAge) +
bstory*(AvgStories) + bYear + errorterm/2)

gen Consumption_adjusted_HA = CA_WOZ/Income
gen Traditional_HA = TR_WOZ/Income
```

```
gen Ratio_Surf = Surface / NS_CA_2
gen Ratio_P = TR_WOZ / CA_WOZ

*Dum_HA
gen Dum_HA = 1
replace Dum_HA = 0 if Ratio_Surf <1
*HA in maanden i.p.v. jaar
gen Traditional_HA_Months = Traditional_HA * 12
gen Consumption_adjusted_HA_Months = Consumption_adjusted_HA * 12

*Equation (5)
gen HA_TR = TR_WOZ / Income
gen HA_CA = CA_WOZ / Income
gen HA_WOZ = WOZ / Income

gen lnHA_TR = log(HA_TR)
gen lnHA_CA = log(HA_CA)
gen lnHA_WOZ = log(HA_WOZ)

*test correlation Ethnicity
corr Country_Born Ethnicity Ethnicity_Generation
drop Ethnicity
*test correlation full model
corr Country_Born Ethnicity_Generation AGE Education Partner Household_Composition Periphery
Year

//reg CA Equation (5) variations R2: 0,4049 - effecten (richtingen) logisch
reg lnHA_CA i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Household_Composition
i.Periphery i.Year i.Province, r
//export
outreg2 using equation5.doc, replace ctitle(Fixed Effects) keep (i.Country_Born i.Ethnicity_Generation
i.AGE i.Education i.Household_Composition) addtext (TFE, Yes, LFE, YES)
display "adjusted R2 = " e(r2_a)

//reg TR Equation (5) variations R2: 0,3730 - coefficienten en richtingen vergelijkbaar met CA
reg lnHA_TR i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Household_Composition
i.Periphery i.Year i.Province, r
//export
outreg2 using equation5.doc, append ctitle(Fixed Effects2) keep (i.Country_Born i.Ethnicity_Generation
i.AGE i.Education i.Household_Composition) addtext (TFE, Yes, LFE, YES)
display "adjusted R2 = " e(r2_a)

//test
reg lnHA_CA Country_Born Ethnicity_Generation AGE Education Household_Composition Periphery
Year Province, r
est store model1

reg lnHA_TR Country_Born Ethnicity_Generation AGE Education Household_Composition Periphery
Year Province, r
est store model2
```

```
suest model1 model2
test [model1_mean = model2_mean]

//test 2
reg lnHA_CA i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Household_Composition
i.Periphery i.Year i.Province
est store model1

reg lnHA_TR i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Household_Composition
i.Periphery i.Year i.Province
est store model2

suest model1 model2
hausman model2 model1, alleqs constant
//test of H0: difference incoefficients not systematic. Chi2(35) = 1082219,47. Prob > chi2 = 0.000.
Therefore, H0 rejected. Models are significantly different from one another.

//Equation (6) - prep
bysort Year: gen P_ca = CA_WOZ / Income
gen lnP_ca = log(P_ca)

*create RHA
bysort Year: egen P_ca_mean = mean(P_ca)
bysort Year: gen RHA = P_ca - P_ca_mean
gen DumHA = 1
replace DumHA = 0 if RHA <= 0
bysort Year: gen RHA_below = P_ca_mean - P_ca if DumHA==0
replace RHA = RHA_below if RHA <=0

*create logs
gen lnRHA = log(RHA)

//MODEL EQUATION (6) //YEARS ONLY
//regression - column 1 - 2
reg lnRHA i.Year i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Household_Composition
i.Periphery i.Province if DumHA==1, r
outreg2 using equation6-year.doc, replace ctitle(Model 1) keep (i.Year) addtext (Demographics, Included,
LFE, Included, Scale, National)

reg lnRHA i.Year i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Household_Composition
i.Periphery i.Province if DumHA==0, r
outreg2 using equation6-year.doc, append ctitle(Model 2) keep (i.Year) addtext (Demographics, Included,
LFE, Included, Scale, National)

//regression - column 3 (severely burdened) - 4 (least burdened)
xtile P_ca_q = P_ca, nq(5)

reg lnRHA i.Year i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Household_Composition
i.Periphery i.Province if P_ca_q==5, r
```

```
outreg2 using equation6-year.doc, append ctitle(Model 3) keep (i.Year) addtext (Demographics, Included, LFE, Included, Scale, National)

reg lnRHA i.Year i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Household_Composition i.Periphery i.Province if P_ca_q==1, r

outreg2 using equation6-year.doc, append ctitle(Model 4) keep (i.Year) addtext (Demographics, Included, LFE, Included, Scale, National)

//test midden quantile
reg lnRHA i.Year i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Partner i.Household_Composition i.Periphery i.Province if P_ca_q==3 & DumHA==0, r
outreg2 using equation6-year.doc, append ctitle(Model 5) keep (i.Year) addtext (Demographics, Included, LFE, Included, Scale, National)

//CHOW TEST
recode Periphery (1=0) (2=0) (3=0) (4=1) (5=1)
label define Periphery 0 "urban" 1 "rural"
label values Periphery Periphery_NL
gen rural = Periphery==1
gen urban = Periphery==0

//Chow-test equation 5
* Pooled model 1 reg
reg normal model , r
*outreg2 using chow_test, replace word dec(4)
* Model 1A reg
reg normal model if rural==1 , r
*outreg2 using chow_test, append word dec(4)
* Model 1B reg
reg normal model urban==1 , r
*outreg2 using chow_test, append word dec(4)

*column 1-2;
xtile P_ca_q = P_ca, nq(5)
reg lnRHA i.Year i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Partner i.Household_Composition i.Province if P_ca_q==5 & rural==1, r
outreg2 using chow.doc, replace ctitle(Model 1) keep (i.Year) addtext (Demographics, Included, LFE, Included)

reg lnRHA i.Year i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Partner i.Household_Composition i.Province if P_ca_q==5 & urban==1, r
outreg2 using chow.doc, append ctitle(Model 2) keep (i.Year) addtext (Demographics, Included, LFE, Included)

*column 3-4
reg lnRHA i.Year i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Partner i.Household_Composition i.Province if P_ca_q==1 & rural==1, r
outreg2 using chow.doc, append ctitle(Model 3) keep (i.Year) addtext (Demographics, Included, LFE, Included)
```

```
reg lnRHA i.Year i.Country_Born i.Ethnicity_Generation i.AGE i.Education i.Partner
i.Household_Composition i.Province if P_ca_q==1 & urban==1, r
outreg2 using chow.doc, append ctitle(Model 4) keep (i.Year) addtext (Demographics, Included, LFE,
Included)
```