How affordable housing makes regions attractive

Evidence from counties in California, United States of America

Jaap Gerben Vellinga[1]

[1] Master Student Economic Geography, Faculty of Spatial Sciences, University of Groningen

Landleven 1, 9747 AD Groningen, The Netherlands.

Student number: S2411016

Email: j.g.vellinga@student.rug.nl

Supervisor: Dr. V.A. Venhorst

January 28, 2019

How affordable housing makes regions attractive

Evidence from counties in California, United States of America

Jaap Gerben Vellinga[2]

Abstract

California has always symbolized opportunity; a reason for many United States citizens to have moved into this state. The high amounts of internal migrants moving into the counties of California contributed to the development of the state's economy into what it is right now. Yet, its knowledge-based economy is still expected to grow, and more than ever California will rely on internal migrants. Not only from the higher skilled, but also from less skilled workers whose demand will increase considerably. However, counties with the strongest labor markets in California became less affordable, and there is evidence that this drives out-migration of people *from* the state. Still, there is no evidence on how housing affordability works on internal migration *into* the state. The aim of this paper was to estimate the influence of housing affordability on internal migration into California counties. Multinomial logit models show significant increased risks of internal migration (compared to no migration) into California counties when housing affordability in counties increases. These results apply to different age-groups and education levels. In addition, two semi-constructed interviews were conducted in order to obtain a wider understanding of the housing affordability mechanism and internal migration in California. The results from the mixed methods used in this paper can be seen as a contribution to the literature about the housing affordability crisis in California that is available at this moment. Understanding the way housing affordability works on internal migration into California, and counteracting on the individual decisions made because of that, could be vital for the state's future economic prosperity.

Keywords: Housing affordability; Internal migration; California

[2] Master Student Economic Geography, Faculty of Spatial Sciences, University of Groningen

Landleven 1, 9747 AD Groningen, The Netherlands.

Student number: S2411016

Email: j.g.vellinga@student.rug.nl

Supervisor: Dr. V.A. Venhorst

January 28, 2019

Contents

Abstract ... **1**

Section 1 – Introduction ... **3**

Section 2 – Literature review ... **6**

2.1 California's population and economy over the years ... 7

2.2 Internal migration: The driving force behind California's economy ... 8

2.3 The components of internal migration ... 10

2.4 Housing affordability and economic location theory ... 12

2.5 Other predictors ... 18

2.6 Conceptual framework ... 19

Section 3 – Research design ... **21**

3.1 Data ... 21

3.2 Statistical methods ... 24

3.3 Mixed methods: Adding qualitative aspects to quantitative data ... 26

Section 4 – Empirical analysis and results ... **27**

4.1 Descriptive results ... 27

4.2 Estimated internal migration differences by housing affordability from multinomial logistic regression models [General and age-specified models] ... 30

4.3 Estimated internal migration differences by housing affordability from multinomial logistic regression models [Education-specified models] ... 37

4.4 Semi-structured interviews ... 41

4.5 Discussion of the results ... 46

Section 5 – Conclusion ... **49**

5.1 Conclusions ... 49

5.2 Recommendations for future research ... 49

Reference list ... **50**

Appendix A – List of counties in the analysis ... 58

Appendix B – Housing affordability index rates per county in the analysis ... 60

Appendix C – Methodology Traditional Housing Affordability Index ... 64

Appendix D – Syntax ... 65

Appendix E – Interview guide ... 80

Section 1 – Introduction

For the United States (further abbreviated as U.S.) it is widely believed that internal migration rates are higher than in other countries. The belief that someone can pick up and move to a location that potentially delivers better opportunities has long been an important part of the American attitude. From a historical perspective Americans were traveling over the Appalachians prior to the Revolutionary war; relocated from the east- to the west coast during the mid-90s; and moved from regions characterized by poverty and low wages to areas that were more prospective: all being characterized as internal migration. The state of California in specific has relied on internal migration from other states since the Gold Rush started halfway the nineteenth century. Since then, the state went through a major transition from a highly industrialized- towards a knowledge-based economy. The state has always symbolized opportunity, and because of that it attracted many domestic migrants. The high amounts of internal migrants moving into the state helped to develop its economy into that of what it is right now. If the state would have been a country at this moment, it would have had one of the largest economies of the world, one that might also be considered as the most innovative (International Monetary Fund, 2017). California is interesting for internal migration studies because, in contrast to many other states, net in-migration in California counties has been positive for most U.S. socio-demographic categories for a long time. In particular among the higher educated, the most affluent, and among younger people (Frey, 1995; U.S. Census Bureau, 2017). Therefore, California can be seen as a vibrant migration state, and with its net population of almost forty million people it is the most populous state in the country. Internal migration can be seen as the movement of people from one residential location to another within a country – that is movement between geographical defined regions, states, or cities (Greenwood, 1997; Molloy, Smith and Wozniak, 2011; Roseman, 1977). Previous research concerning internal migration and residential mobility provides a great deal of information about how a wide range of variables related to family contexts, life course events, and economic circumstances can generate residential change (Greenwood, 1975; Molloy, Smith and Wozniak, 2011; Mulder and Wagner, 1998; Whisler et al., 2008; and many more). Early research mainly saw residential change as a one-off action instead of a more sophisticated decision-making process involving many different factors. In the more recent literature, however, more emphasis is being placed on obtaining a more sophisticated individual background on why people migrate and the underlying mechanisms. For example, Clark and Lisowski (2017) recently conducted research into the process prior to move and how this process can lead to the actual move.

Internal migration has always been important for California's economy, but for future economic prosperity it might be even more important than ever before (Johnson, 2018). Despite the fact that California possesses some remarkable qualities, it also faces some major problems that are likely soon to weaken its economy. Among these problems are increasing income inequality, labor shortages, and potential rises in ethno racial tension. The major problem for California, however, is the problem of housing unaffordability. Ignoring the problem related to housing unaffordability not only allows it to worsen, it also accelerates the downward spiral because it aggravates over time. There was already evidence that the state is losing lesser skilled people, but recently it also seems that is loses some of the higher skilled workers, who move in patterns that imply that housing affordability is a driving force behind these moves (Bean, Brown and Pullés, 2018; Johnson, 2018; Woetzel et al., 2016). However, if housing affordability is a driving factor behind out-migration patterns *from* California, then such a pattern would also be expected *into* and *within* the state because housing affordability rates highly vary between the counties of the state, ranging from 7% in Santa Barbara to 77% in San Bernardino between 2000 and 2018 (Appendix B; California Association of Realtors, 2018a). The California Association of Realtors (2018a) quarterly computes a housing affordability index for every county in California that measures the percentage of households who can afford to purchase a median priced home. Figure 1 shows that the affordability of housing in California has been declining since the recovery from the financial crisis in 2012. Several studies have looked into the influence of affordable housing on out-migration patterns *from* California (e.g. Bean, Brown and Pullés, 2018; Gunderson and Sorenson, 2010; Woetzel et al., 2016), yet there is hardly any evidence about how housing affordability influences internal migration *into* the state. Understanding the role that housing affordability has on internal migration in California will be useful for either facilitating or

counteracting on residential choice. Undesirable mobility can cause or worsen regional problems, and can obstruct future economic development. Therefore, Coulter et al. (2011) and Clark and Lisowski (2017) argue that understanding the process that causes internal migration is an important contribution to the broad framework of residential change. However, the latter ones were rather skeptical about the influence of places on decision-making related to migration, and argue that place and satisfaction with place, are context-setting but only play a marginal role in the decision-making process. In an article about migration patterns of higher educated people in the larger cities of The Netherlands, Venhorst et al. (2011) argued that regional stakeholders sometimes focus too much on migration of people, while their regional economy actually performs well. In this research, cost of living only has a weak influence on migration patterns; instead it is the region's economic situation that determines migration of young talent (Venhorst et al., 2011). For the U.S., Molloy, Smith and Wozniak (2011) argued – in their article about internal migration – that changes in the housing market contributed relatively little to the internal migration rates within the country until 2011. However, they argue that understanding regional influences on internal migration in the U.S. are clearly an important subject for future work. On the contrary, Herzog and Schlottmann (1986) found remarkable results regarding the spatial-demographic setting of internal migration by using data from the 1980 U.S. Census and the 1981 Places Rated Almenac. They estimated a logit model of leaving versus staying in a region, and found that regional migration rates in metropolitan areas significantly decreased with high cost of living. Whisler et al. (2008) came with an updated research by using the 2000 census and the 1997 Places Rated Almanac. They found similar results related to migration and the cost of living, but found that cost of living either influences the migration of the very young or the very old in U.S. regions. There is a growing interest in studying the migration preferences of different age groups, especially the young adults aged 25-34. They are the ones who carry the future of a place but are also those who strongly respond on factors that influence residential change (Benetsky and Fields, 2015; Franklin, 2003; Goworowska and Gardner, 2012).

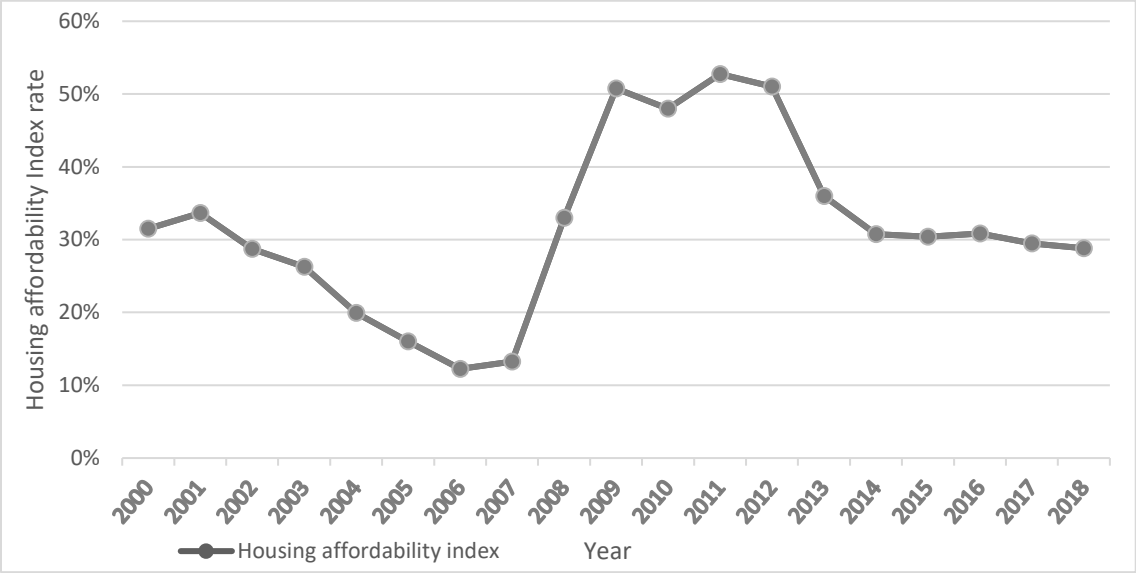
Individual decision making related to migration usually involves two components: The decision *to migrate* and the decision *where to migrate* (Roseman, 1977). In California, there is a tendency to live in the urban counties, since these are the places with the highest quality amenities, including the strongest high-tech labour markets. However, according to Alonso's Theory of Urban Land Market (1964) this drives up the competition for (urban) land with the consequence that houses got less affordable in these areas. The study in this paper also emphasized on how the influence of housing affordability on internal migration into California was different between age- and education groups, because it was expected that California's knowledge-based economy will heavily rely on younger adults, and on people with different skill levels – thus not only the higher skilled (Johnson, 2018). The study engages on the question about the correlates for housing affordability rates on internal immigration patterns in counties in California, and these correlates for different age- and education groups. In the statistical analysis there has been modeled at two levels: 1) All individuals who either move or not move into California counties, and 2) individuals in specific age- and education groups who either move or not move into California counties. For both models the dependent variable is represented by three categories: Not moving, moved from within the same county, and moved from another county. The analysis has been performed by using cross-sectional individual microdata data from the Annual Social and Economic Supplement (ASEC) compiled by Flood et al. (2018). It should be emphasized that the ASEC only provides publicly available individual migration data about the destination county – that is the county where the respondent currently resides in and thus the county where the respondent moved to. Therefore, this study only focuses on internal migrants who move within or into a county due to these data limitations. Thus, this paper only focuses on destinations of internal migrants and examines their individual decision to migrate into California counties but does not study their origins. One of the consequences of these data limitations is that the housing affordability rates from origin counties cannot be compared with those of destination counties if respondents moved between the boundaries of counties. Therefore, this paper does not aim to provide evidence about whether housing affordability can be seen as a push factor influencing the individual's decision *to migrate*, instead it examines whether regional housing affordability is a pull factor influencing the individual's decision *where to migrate.*

To elaborate on the development of recent literature that tries to obtain more knowledge about the underlying mechanisms that influence migration choice by individuals, this paper aimed at finding empirical evidence for one particular mechanism in California, namely regional housing affordability. In specific, the goal of this research was to estimate the influence of housing affordability on internal migration into California counties, deriving either from people who moved from the same county or those who came from other counties in California and elsewhere in the U.S. Although the emphasis of the study in this paper was placed on quantitative research by providing an estimation model of internal migration by housing affordability, the relevance of qualitative research methods cannot be neglected. Therefore, two semi-constructed interviews were conducted with two researchers who had either academic or business/economic interests. Combining quantitative with qualitative research is also referred to as the methodology of mixed methods (Winchester, 1999). In order to understand the relationship between housing affordability and internal migration into California counties, the following main research question and supplemental sub-questions were prepared:

> Main question: *What is the influence of housing affordability on internal migration into California counties?*
>
> Sub-question 1: *To what extent does housing affordability in counties influence the decision to migrate into California counties?* [Quantitative]
>
> Sub-question 2: *How is the extent to which housing affordability influences internal migration into California counties different between age groups?* [Quantitative]
>
> Sub-question 3: *How is the extent to which housing affordability influences internal migration into California counties different between education levels?* [Quantitative]
>
> Sub-question 4: *How does housing affordability influence internal migration into California counties?* [Qualitative]

The main question asks whether housing affordability has influence on internal migration patterns in California. Can there be elaborated on the work of Whisler et al. (2008) who claim that regional context matters on residential choice? Do – as Clark and Lisowski (2017) argue about place and satisfaction with place – California counties with regard to affordable housing only marginally influence the migration decision? Or does housing affordability not have any influence on internal migration into California at all? In order to answer the main question, four sub-questions were prepared. The first sub-question relates to the extent in which housing affordability influences internal migration into California counties, and therefore this question closely relates to the estimation of internal migration by housing affordability. On the socio-demographic front two important questions concerning age and education levels were asked for which housing affordability may be a factor of internal migration. Even though this paper was designed to estimate internal migration by housing affordability and relies mostly on quantitative data, it is also useful to understand the wider context in which this happens. Therefore, the fourth and last sub-question asks how affordable housing can have influence on internal migration into California counties. As mentioned before, two semi-constructed interviews were conducted to answer this question. By putting it all into a nutshell, the objective of this paper was to answer the research questions as shown above and, thus, to reveal whether there is a relationship between regional housing affordability and internal migration into California counties. The next section provides a literature review and elaborates on earlier research and theories related to internal migration and housing affordability. It also extensively describes internal migration and economic changes in California throughout the years, which in turn have a major impact on housing affordability in some counties. The other sections in this paper elaborate further on the research questions as prepared above.

Figure 1 Housing affordability in California (2000-2018)



Source: Own elaboration on data by California Association of Realtors (2018a)

Section 2 – Literature review

The U.S. has long been known by having way higher migration propensities than other countries. Migration and natural change – that is births and deaths – are components of population change. Migration can be seen as the movement of people from one residential location to another (Roseman, 1977). A broad definition characterizes migration as a "move over long-enough distance to entail an appreciable change in the local economic environment" (Molloy, Smith and Wozniak 2011, p.174). By examining population change within a country, migration is usually a more influencing component than natural change. Within the U.S., for example, the net population variation from migration is much larger than the shift from births and deaths. In his resource paper about changing migration patterns within the U.S., Roseman (1977) states that decision-making related to migration is not only based upon the choice to change a specific residential environment, but it is also a decision to relocate the "home-base" for the household activity space – that is the set of places in which the household interacts on a daily basis in terms of work, shopping, socializing and education. Therefore, the decision represents both a change in household and its relative location. The development of new data sources allowed researchers to define migrants and answer research questions about migrants more precisely. As migration is a component of population change, internal migration – sometimes referred to as domestic migration – can be seen as a component of migration. Internal migration can be seen as human population flows between geographical boundaries within a country, which are regions, states, or cities (Molloy, Smith and Wozniak, 2011). In 1997, Greenwood was able to analyze the equilibrating effects of internal migrants on local economies in the U.S. and provided a comprehensive literature overview. Complementary, Molloy, Smith and Wozniak (2011) added fifteen more years of data to Greenwood's overview and analyzed internal migration in the U.S. from 1980 to 2009. However, little academic literature is available about internal migration patterns from the recent housing market contraction and the recession until now. To define migrants, one should consider both the geographic units to define potential origin - and destination locations and the time period in which individuals must move between origins and destinations (Long, 1988, p. 4-12; Molloy, Smith and Wozniak, 2011). Traditionally, migration students made a distinction between local movers and migrants, where local movers are those who moved within the boundaries of the county while migrants crossed the county borders (Roseman, 1977). In this paper local movers and migrants were both seen as internal migrants. A county is a geographical region of a country used for administrative or other purposes (Chambers Dictionary, 2005). Within the U.S. there are 3,242 counties and the state of California counts fifty-eight of them. Traditional approaches to study migration tend to use

aggregate ideas and concepts, while more recent studies tend to focus more on the individual decision-making of migrants. This paper elaborates on those recent studies by focusing on the individual's decision to migrate because of housing affordability. Roseman (1977) has subdivided the individual's decision to migrate into two categories: The decision *to move* and the decision *where to move*. Affordable housing can be seen as a factor that may influence the individual's decision *where to move,* although it could also apply to the decision *to move.* In this paper, housing affordability is defined as the percentage of households that can afford to purchase a median priced home in California counties based on traditional assumptions (Appendix C; California Association of Realtors, 2018b), and therefore it best fits the individual's decision *where to move*. Nevertheless, both the decision *to move* and the decision *where to move* are elaborated in this section in order to better understand the internal migration mechanism. Before elaborating on these, it is useful to understand the context in which internal migration was studied. As mentioned in the introduction of this paper, there was chosen to focus on the geographical scale of counties in California. Therefore, this section starts with an overview of California's population and economy in order to better understand the circumstances in which housing affordability may influence the state's internal migration.

2.1 California's population and economy over the years

Internal migration rates into California have always been higher than any other U.S. state. According to Starr (2007) the state has symbolized opportunity since the Gold Rush that started in the mid-nineteenth century, a characteristic that boosted California's popularity among U.S. residents. The state has the third largest area after Alaska and Texas. Even more notable are its population and economy. As mentioned in the introduction, the state has more than forty million inhabitants, which makes it also the most populous U.S. state. Yet, its population is highly concentrated and not evenly distributed. By looking into the state's economy, it produced a gross domestic product of almost $ 2.8 trillion in 2017. This makes California the state with the largest economic output in the U.S, and even more remarkable one of the largest economies in the world if it would have been a country[3] (International Monetary Fund, 2017; US Bureau of Economic Analysis, 2017; Woetzel et al., 2016). However, even though the 1849 Gold Rush and some other huge state investments – such as the construction of intercontinental railroads – objected to be pull-factors for internal migration, California's population did not really increase with the expected amounts of people until the Second World War. In their overview about internal and international migration in California, Bean, Brown and Pullés (2018) distinguished three unique periods of developments in California, namely *the early development phase* (1900-1940), *the industrial boom* (1940-1990) and the *post-industrial, high tech era* (1990-current).

The *early development phase* started in the beginning of the twentieth century, and during this period California's population growth was a little higher compared to the U.S. as a whole, due to westward moving internal migrants and somewhat to international migration. As mentioned before, California symbolized opportunity, which worked as a magnet on internal migrants. However, it was the *Industrial boom* where the state's population really started to grow. During this period World War II came to an end, but U.S. engagements in the Cold War, Korea and Vietnam just started to happen. Because of this, there were – especially in California – massive investments in shipbuilding, defense and the aerospace industry (Hersch, 2015 p.30). This has led to a huge growth in the state's manufacturing industries. Large-scale agriculture developments in the central counties of California also largely contributed to the state's well working economy. The huge investments in extensive agriculture even pay off today, since California is still the number one food provider of fruits, vegetables and nuts of all U.S. states (U.S. Department of Agriculture, 2017). The state's flourishing economy in combination with large developments in infrastructure, including massive public investments in higher education, transportation (primarily in roads) and water projects (Cohen, 2011; Starr, 2005) made the state very attractive for internal migration. California experienced a

[3] Calculated by comparing state-level GDP from the U.S. Bureau of Economic Analysis (2017) with global data from the International Monetary Fund (2017).

disproportionate population growth during this period. Yet, at the beginning of the *post-industrial, high tech* era there was break in the rapidly growing population and economy. Because of there being no engagements in Korea and Vietnam anymore and because the Soviet Union collapsed, the national government started to spend less on aerospace and armory. California in particular was hit by decreased governmental spending in these industries. A three-year recession was the result in which the gross domestic product (GDP) fell substantially, many people lost their jobs, and when the population growth declined to approximately the national average (Bean, Brown and Pullés, 2018). People got more and more concerned because of the poor economic situation, but fortunately the effects of large-scale investments in infrastructure and higher education gradually became visible. High-tech companies in Silicon Valley and other parts of the state began to flourish, and created a foundation for future economic prosperity. Many of the unemployed manufacturing workers found jobs related to the production and assembly of computers and chips (Rhode, 2001). As a result, California's traditional economy that relied on manufacturing and agriculture started to transform towards a high-tech, knowledge-based economy. The productivity of these high-tech industries provided a major boost to the state's economy, which still – as of today – is a major economic driver for California. However, the transition from the industrial periods to the post-industrial, high-tech era also implied changing internal migration patterns into the state. In 1990, the massive population growth dropped and only about 40,000 more internal migrants entered the state than left. Besides, about eighty-five per cent of those who migrated into the state acquired at least a Bachelor's degree – or four year college degree – or higher, reflecting the state's high-tech evolution. The population growth declined more in subsequent years, and 2000 was the first year where the net number of internal migrants had turned negative. Meanwhile the domestic population growth of the higher skilled in California kept increasing while the state was losing the lesser skilled ones (Bean, Brown and Pullés, 2018; U.S. Bureau of the Census microdata from Ruggles et al., 2017). In their paper, Bean, Brown and Pullés (2018) defined the higher skilled as those who acquired a bachelor's degree or higher, while the lesser skilled are referred to as the ones who only completed high school or less. So, it can be argued that the economic transition involved a notable internal migration change. There where California used to gain lesser-skilled migrants during the industrial eras, it is now mainly attracting the higher skilled people in its knowledge-based economy. Even though California is still seen as a well-positioned state with the potential of future economic growth driven by its new industries (U.S. Bureau of Labor Statistics, 2018; Vara, 2015), it should be emphasized that also the gains of net higher skilled internal migrants are now slowing. According to Bean, Brown and Pullés (2018) the post-industrial, high tech period most recently involves the beginnings of out-migration of both lesser-skilled and some higher-skilled workers who move in a pattern that implies that housing affordability is the driving force behind these moves. But why is internal migration, or domestic migration, important for the state of California at this moment? The economic importance of internal migration into California was elaborately explained in the following sub-section – in specific the importance of internal migration of younger workforce talents and various skilled workers into the state.

2.2 Internal migration: The driving force behind California's economy

California has long been known for its overwhelming population growth. No other developed region in the world has ever experienced and sustained the magnitude and duration of population growth as California has. Before the post-industrial, high-tech era, most of California's population growth came from internal migration, while nowadays the population growth in California is mainly produced by natural increase. Nevertheless, in his report about the future population of California, Johnson (2018) projects continued population growth for California, not by internal migration though but by natural increase. So, if population growth is expected to continue to grow, even though it was not by gaining migrants from elsewhere in the U.S., why – then – is internal migration so important for California? The answer to this question lies in the labor market of the state. The projected continuation of population growth for the next decades is mainly caused by natural increase. For California this implies that the population is aging, which potentially may harm its robust workforce (aged 20-64). Large numbers of Californians are reaching retirement age or already achieved that age,

and in the next decade the majority of residents will be in those age groups. Data from the California Department of Finance projections (Johnson, 2018) shows that throughout years the relative portion of Californians with ages of sixty-five and over increased from 9.0 per cent (1970) to 13.6 per cent (2016) and is objected to grow to 19.0 per cent by 2030. Meanwhile the number of children will increase very slowly due to declines in birth rates and small increases in the number of women of childbearing age. The state's changing population will put pressure on infrastructure, public services and on the state's economy as a whole. These include education, transportation, housing water, food, and healthcare sectors. Since it is unlikely that California's natural population is able to meet the demand of the future labor market, people with younger ages have to come from elsewhere. Therefore, the demand for internal migration of younger workforce talents into California's counties – especially those counties with the strongest labor markets – remains high in the nearby future. There is a growing interest in studying the migration patterns of the young adults. From a historical perspective the young adults, aged between twenty-five and thirty-four, have always been more mobile than the rest of the population (Benetsky and Fields, 2015; Franklin, 2003; Goworowska and Gardner, 2012; Kodrzycki, 2001; Saks and Wozniak, 2011), while internal migration propensities fall with age (Molloy, Smith and Wozniak, 2011; Roseman, 1977) According to Roseman (1977) the young adults are the ones who experience most key changes in their life- and career cycles, and therefore their internal migration propensities are higher than other age groups. More information about these cycles was provided later in this section. It is the young adults who are usually known for their large contribution to the share of internal migrants in the U.S. (Benetsky and Fields, 2015; Franklin, 2003; Goworowska and Gardner, 2012). In her study to the migration of the young adults between 1995 and 2000, Franklin (2003) argues that their migration choices may be influenced by local amenities, including housing preferences. Understanding the migration patterns of this group of migrants and how these differ from the rest of the population may provide important insights into their location preferences. Maintaining or attracting young age groups are important policy objectives for most regions, anywhere (Goworowska and Gardner, 2012; Whisler et al., 2008). In addition to the fact that they may potentially provide population growth through future childbearing, an increasing share of the young adults are highly educated, which "provides a measure of economic opportunity in the area, while simultaneously serving to raise the stock of human capital" (Franklin, 2003 p.2). An increased level of human capital then ensures future economic growth in sectors where education has a major role. With fast developing high-technology industries in many California counties, but also with increased demand in the less- and middle skilled job markets and the aging population of the state, it can be stated that either keeping or attracting younger workforce talents into the state's counties is essential to secure a sustainable economic future.

It has become clear that California's economy has relied both on the lesser- and higher educated internal migrants from the beginnings of the 20th century until now. The state's future economic situation is also expected to rely on both facets. Higher educated people are necessary to ease up the demand for labour in the higher-skilled industries, while less and middle-educated workers could release the pressure on infrastructure and public services (Johnson, 2018). Education, sometimes referred to as human capital, has a major role in the economic development of nations, regions and cities (e.g. Barro, 1991; Lucas, 1988; Glaeser and Mare, 2001). According to Florida (2017, p.103) education "is a key factor in how much money we make and reinforces and reproduces the advantages that money brings". Three different education groups were defined according to Bean, Brown and Pullés (2018) and Fuller and Raman (2016). Those who were referred to as lesser educated did not complete anything higher than high-school, while the higher-educated individuals were the ones who graduated with at least a bachelor's degree. Lesser educated people have to deal with considerable issues in society because they earn way less than people with higher levels of education. Besides, they often experience substantial higher levels of unemployment. The higher educated – in contrast – usually have incomes and salaries much higher than the lower educated while they also experience way lower unemployment levels. Therefore, they are advantaged with numerous economic benefits compared to the lower educated groups (Bean, Brown and Pullés, 2018; Florida 2017, p.103). In addition to the lesser- and higher educated, individuals who were middle educated were also included in this research, since they may be the ones whose services are most needed in the future economy (U.S. Bureau of Labor Statistics, 2017; Woetzel et al., 2016). The middle-skilled jobs require more

education and training than high school but less than a four year college degree - or bachelor's degree (Fuller and Raman, 2016). It can be argued that the state's current economic situation heavily relies on sophisticated and innovative industries, and therefore depends on the higher educated people coming from within the state, other states, and other countries. However, it should not be forgotten that also the less and middle educated workers are important for California to sustain its current knowledge based economy. According to the U.S. Bureau of Labor Statistics (2017) the knowledge based, high-tech economy will continue to grow, which also implies a higher demand for service- and construction workers, including food service and health-care. How education works on residential decision-making is actually quite simple. Members of higher educated classes have more money; are able to bid higher prices for land; and therefore have the greatest ability to choose where to live. Section 2.4 elaborates more on how the affluent individuals are able to bid higher prices for land than the less moneyed ones, which gives them the opportunity to find residence at the most desirable places – also known by the "Theory of Urban Land Market".

Because the state has always relied on domestic (and international) migration, Bean, Brown and Pullés (2018) expect that there is little reason to think the future economic situation will not require these same kinds of migration. However, they are concerned that there are threats that may obstruct migration into California, including the threat of housing affordability or rather housing unaffordability. In their research they compared recent population growth of California with Texas from 2016 to 2017, another large state with a dynamic economy and often seen as the number one competitor of California. A remarkable finding was that Texas gained about 80,000 domestic migrants on a net basis while California lost almost 140,000 to other states that year. Factors such as economic weakness, low demand on the labor market, or high supply from elsewhere do not explain this pattern for California. Thus, the negative net domestic migration must be explained by some other factor. Bean, Brown and Pullés (2018) think that the factor most likely explaining this pattern is California's extremely high housing costs, which are driving the working-class and middle-class people out of the state. Johnson's report (2018) supports their claim by showing that the state indeed is losing the lower educated Californians while it gains those with a college degree or more. If housing costs are the factor behind out-migration, it is expected that the lesser skilled are more affected than the higher skilled. Other evidence comes from Gunderson and Sorenson (2010). By using simple and augmented gravity models the authors do not find evidence of out-migration patterns from California counties by the state's economic situation. However, they found that housing affordability had an important influence on how California's traditional attraction for internal migration has reversed. Nevertheless, it must be emphasized that also within the state the housing costs highly vary between its counties (California Association of Realtors, 2018a). So, if housing affordability drives people out of the state, then an internal migration pattern within the state by housing affordability would also be expected. But how do houses get unaffordable in some counties, and why is housing affordability so important for a region's economic sustainability? The next subsection focuses on the individual components of internal migration, namely the decision *to migrate* and the decision *where to migrate.* In addition, to explain the disproportionate housing unaffordability in some of California's counties, Alonso's Theory of Urban Land Market (1964) was elaborated in the section 2.4.

2.3 The components of internal migration

From the previous sub-sections it has become clear that internal migration into California was and will be important for the state's economic development. Internal migration can be seen as a component of population change, a component that results in a major adjustment: Residential change. The individual's decision to internally migrate is often a result of two components: A set of individual decisions *to move*, and a consequence of decisions *where to move* (Roseman, 1977). Even though it can be difficult to distinguish between the two decisions, which are sometimes made simultaneously, separation of the components contributes to understanding the individual internal migration mechanism.

The decision to move. The decision *to move* is often a very ordinary and expected part of life. According to Roseman (1977) decisions to move are related to two cycles: The life cycle and the

career cycle. Life cycle events can relate to various natural adjustments. As individuals often leave their parental home upon graduation from high school, they usually form a single-person household. Subsequently, marriage usually results in one or two moves. As a follow-up, family expansion often involves the need to change housing, again leading up to a decision to move. American middle- and upper-middle-class young couples are characterized to move from an apartment to a single house when they get children, and change to an even larger house when the family further expands. At the end of the life cycle, children grow up and older couples usually move into smaller houses as a result of this key change. Separations, divorces and other family changes may also lead to additional moves. Certain key points in the career cycle also influence the decision to move. The acquisition of a job after college or high school, job transfers, lay-offs, dismissals, or promotions can directly be related to the decision to migrate. Job promotions can also indirectly influence the decision-making process by providing an extra boost of financial capital. Besides, people often tend to match their residential house to the status of their job. Once individuals retire they are not bound to the activity space related to working, and the propensity to migrate increases as they want more leisure time. Other decisions to move have little to do with life- or career cycle changes. Some decisions are forced due to – for instance – infrastructure construction or nuisance. In fact, the decisions to move are in these cases not made by migrants, but rather by governmental and/or institutional bodies (Stark and Taylor, 1991). Other decisions to move can be related to neighborhood conditions, such as threats to property values, safety, and school quality (Greenwood, 1975; Lee, Oropesa and Kanan, 1994). In the statistical analysis performed for this research, a series of controls are taken into account that relate to the decision to migrate. These are extensively elaborated in section 2.5.

The decision where to move. While the decision *to* move primarily depends on a series of personal characteristics related to life- and career cycle events or to external factors, the decision *where to move* mainly depends on the characteristics of the destination place. It is, however, difficult to predict migration for particular places. Early approaches for examination attempted to predict migration by labor market conditions, such as wage levels and composition of the labor force. This method, however, suffers to some constraints. By examining migration among major metropolitan areas in the U.S., Lowry (1966) found that places with more favorable economic and job climates got more in-migrants. However, there were no significant differences found related to out-migration between more and less favorable places in terms of labor market conditions. Apparently only in-migrants responded to these economic conditions. This theory is also known by the Lowry Hypothesis. From an older research perspective, households were indifferent between attractive and unattractive places. Since location-specific amenities both involve labor and house-markets, attractive places implied high rents and low wages, while unattractive places entailed low rents and high wages. According to Bloomquist, Berger, and Hoehn (1988) this principle worked both at inter- and intra-regional levels. In his study about changing migration patterns in the U.S., Roseman (1977) states that it is most likely that people move again right after the period when they have moved. So that would contribute to being indifferent between attractive and unattractive places. However, Roseman also found that the probability of moving declines with time at a given place; as people establish social and economic ancestry at a particular place they tend to stay. The decision to move again disappears when people when people become more familiar with the place and create loyalties. And, that is why location matters. In a recent study about migration and the quality of life, Whisler et al (2008) found that places with wide varieties of amenities are more favorable than those without. Amenities can be defined as "site- or region-specific goods and services that make localities more or less attractive to agents" (Whisler et al., 2008 p.60). These can – for example – relate to availability of consumer goods and services, entertainment and recreation, education, and employment. Recent research from among other Glaeser (1998) and Florida (2002) shows that an increasing number of households, especially the young-educated, have strong desires to live in certain kinds of urban places to fulfill their lifestyle demands and maximize their utility. According to Whisler et al. (2008) urban places contain a wide range of consumer goods and services, diverse entertainment and recreational opportunities, dense networks of education, employment, and social opportunities, and tolerant racial and social attitudes. Moreover, urban areas have a large share of the world's most successful industries, innovation, start-ups and high-skilled people (Florida, 2017 p.6; Florida & King, 2016). Therefore, one could argue that the cost of living in these places would not matter because people aspire to live at those places where the most ambitious

and talented people want to be; those places with a wide variety of amenities. Besides, growing tax revenues from high skilled workers in urban areas lead to higher developed schools, transit, libraries, parks, and so on. Higher quality amenities, therefore, increase the advantages of those who live close to urbanization. In his article about the richest cities for young people, Thompson (2015) refers to this as the ongoing feedback loop that ensures the success of these places over time. Besides, most high tech industries are located in the urban coastal counties of California, while the more vulnerable agriculture and manufacturing sectors were based in the periphery (Glasmeier, 2015 p.1 – 23), making these urban places even more attractive than the periphery. Place specific amenities have an important influence on internal in-migration patterns, and urban milieus have a wide variety of these amenities. Because of that, urban areas have pushed together talented people with different race and ethnicity from all over the world, which also happened in the urban counties of California. This process can be seen as clustering of people into urban areas, but according to Florida (2017, p.21) this "self-reinforcing process generates its own fundamental contradiction." Even though clustering drives growth, it also drives competition for limited urban space, which in turn drives up the housing prices. The way in which competition for land works on housing affordability within counties is explained in the next sub-section.

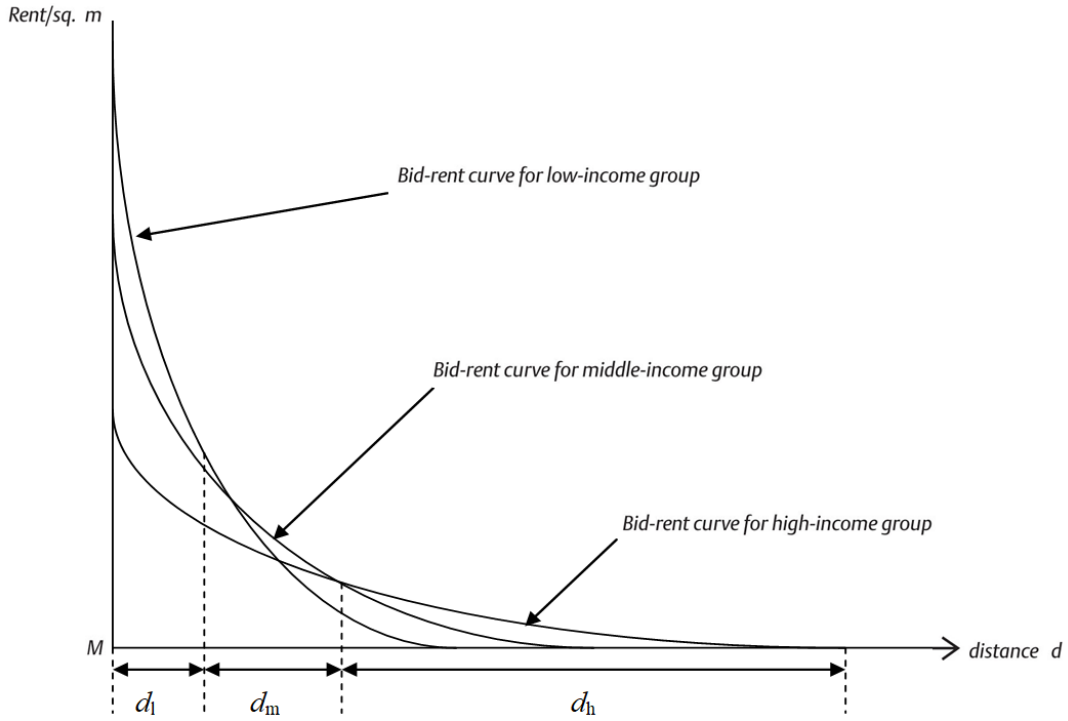
2.4 Housing affordability and economic location theory

The importance of location is often not recognized when policymakers discuss about housing. Rather, they try to focus on supply or the total number of subsidized units, where location is just being seen as another amenity such as the presence of an extra bedroom (Pickford, 2016). Understanding housing affordability in a region is more than just looking at supply- and demand in the area. Even though supply- and demand calculations contribute to the process of understanding housing affordability, they do not provide any insight in the relationship between price and location, which is of great importance for understanding the housing affordability mechanism. As mentioned before, housing affordability is defined as the percentage of people that can afford to purchase a median priced home in California counties. Therefore, housing affordability is both a measure of household income and the median price of homes in California counties (Appendix C; California Association of Realtors, 2018b). The median household income in California rose from about $47,000 dollars in 2000 to almost $70,000 in 2017 (U.S. Census Bureau 2000; 2017). Yet, the counties where median household incomes increased the most – that is the highly concentrated urban counties – were also the ones where houses became least affordable. From that perspective, it can be argued that housing prices did not rise correspondently with the incomes of households in these counties (Quigley and Raphael, 2004). The relationship between house prices and location in California counties can better be explained by economic location theories including Alonso's Theory of Urban Land Market (1964). In contrast to firms that determine their optimum location by economies of scale and agglomeration economies – given their product or service, production technology, customers and suppliers – individuals choose their location by finding the place that satisfies their space needs and location preferences with their budget constraint (Alonso, 1964; Dieleman & Wegener, 2004; Von Thünen, 1826). One of the early pioneers in economic location theory was J.H. von Thünen (1826), who implied that the price (or rent) of any place (or location) equals the value of the product minus production and transport costs. Under the assumption that transportation costs for any product are the same, activities with the highest production costs are located near the marketplace while those with lower costs related to production will find their location further away. Inspired by von Thünen's theory, Alonso (1964) created a model for both firms and households, also known by the "Theory of Urban Land Market". While Von Thünen applied his model to the cost of agricultural land, Alonso extended his theory to modern cities and residential land values. The theory simply but clearly describes the competition for space, where the price of land follows a series of bid rent curves in which land prices decrease when it is located further away from the most competitive and desirable centers. Bid rent curves explain how location prices, or the value of a place, are determined. For households, who – in contrast to firms – do not have any cost functions, the locational trade-off is between land consumption and the distance to the urban center. The expenses of an individual household related to locational choice, thus, depend on their division between land and transportation

costs. By elaborating on von Thünen's model, Alonso assumed that prices would be the highest in cities, due to lower transportation costs. Households try to maximize their combined utility of land and transport given their budget constraint. Historically, this implied that high income households occupied large sites at the periphery to avoid congested, noisy, and dirty districts. On the contrary, low-and middle-income households chose their location in high-density housing areas near the urban center because they could not afford to exchange land for other factors and did not have the ability to pay for the costs to get to the city (McCann, 2013 p.123). Figure 2a shows how this mechanism worked. First, the competition for housing is between all income groups where every group has a different bid-rent curve. The figure shows how the slopes for each bidder are different. The slopes indicate the amount that bidders are able to pay as the distance from the urban center increases. It becomes clear that high-income groups were able to exchange smaller distances to the urban centers for land consumption and higher transportation costs, while low-income groups were bound to the urban labor markets and could not afford to live in more rural areas. The bid-rent curves of the low-income groups are steeper because the transport costs related to increasing commuting distance quickly reduce the money they have available to spend on land and non-land inputs (McCann, 2013 p.124). The land occupied by the high-income groups is between M and d_h, while the land occupied by the lower income groups was respectively between M and d_m (middle-income) and M and d_L (low-income). To be able to co-exist with each other, the slopes of the bid-rent curves of the three groups must be different. That is why the slopes of the higher income groups are shallower than those of the lower incomes. Another important assumption in this model is that when income increases, individuals have higher preferences for land consumption, which is stronger than any preference for increased accessibility to the city center. The land allocation results from the figure were based on the strong assumptions related to the behavior and preferences of the different income groups (McCann, 2013 p.123).

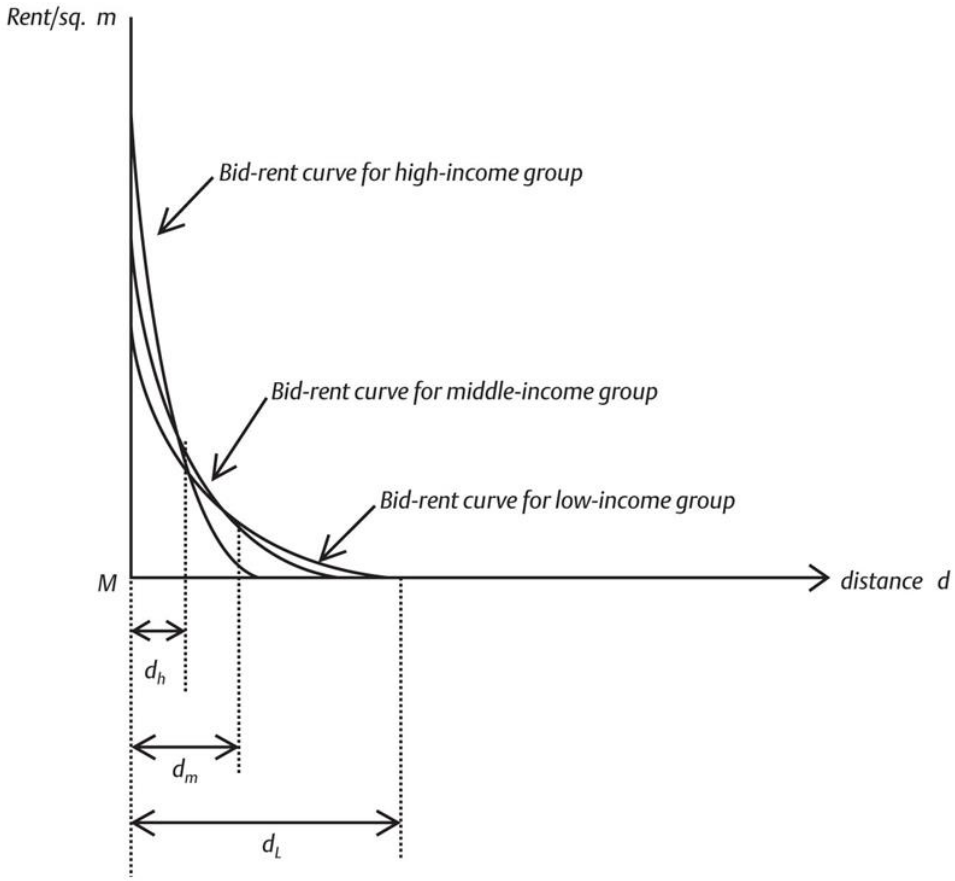
Between 1980 and 2000 most urban areas in California consisted mainly of poor and black people. During this time rich and slightly less rich people left the urban areas. This all changed from 2000 when the richer people retreated to the urban centers. "The back-to-the-city movement" – as cited by Florida (2017, p.58) – has had some consequences including increased housing prices in these areas, falling most heavily on the poor and disadvantaged. This is a trend not only going on in California but also in the U.S. as a whole as argued by Baum-Snow and Hartley (2016) in their working report about the transformation of downtown neighborhoods in 120 US metros between 1970 and 2010. Between 2000 and 2010, the share of higher income and higher educated white households living in urbanization increased in almost 70% of the studied metros. As mentioned before, urban areas offer access to amenities – from libraries and monuments to cafés and theatres. This is often seen as the most important factor driving the affluent ones back to the urban areas nowadays. Additionally, other factors such as a large concentration of high paying and creative jobs and reduced commutes to their jobs also contribute to the back-to-the-city movement of the more affluent people (Florida 2017, p.64). Yet, one of the negative consequences of this movement is the great demand for housing in the counties possessed with these advantages. The fierce competition for urban space drives house prices up, and as a result the less affluent ones are being prized out from these areas. Figure 2b shows how this works in economic location theory. Again the distance between d_h, d_m, d_L and M represents the land that each group occupies. In contrast to the historical perspective, the assumption regarding the relative preferences for space and accessibility are not justified now, since it is more attractive for the high-income groups to live in the city center. Therefore, the fundamental assumption in figure 2b is that highly accessible locations are more attractive and of higher value on the housing market than the periphery (McCann, 2013 p.126). As a result the bid-rent curves get steeper and those with higher incomes price out lower income groups. The less affluent households simply cannot afford to pay for a living anymore. Besides, as more people move to a specific location, that location becomes more valuable for commercial users as well; this increases the competition even more (Florida and Mellander, 2016).

Figure 2a Residential urban land allocation for different income groups [traditional]



Source: McCann (2013 p.123)

Figure 2b Residential land allocation with high relative preferences for accessibility [recent]

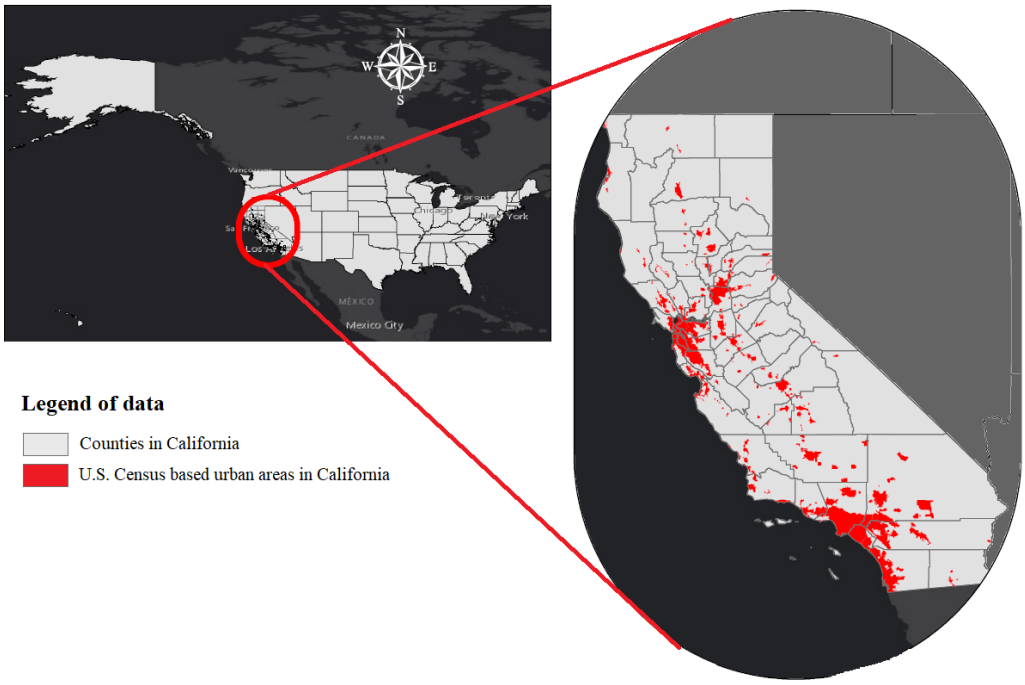


Source: McCann (2013 p.126)

As of today, affluent people are taking over the urban areas, once occupied by the lower income groups and industry and commerce, while bidding up the prices. If more people want to cluster in (urban) space; the more competition there will be; the more expensive land will get; and the higher house prices will become, eventually making houses less affordable. And that is exactly what Florida (2017) describes in his book about the new urban crisis, where he points out that technicians, professionals and other high-income groups moved back to urban cores while the less advantaged are being priced out. On the one hand, Florida (2017, p.8) describes that clustering of talented and ambitious people in urban areas is "the basic engine of innovation and economic growth." On the other, as also described before, future economies also rely on the lesser- and middle-income groups who carry a large share of future economic prosperity. Alonso's theory of Urban Land Market helps to understand that affordable housing is not just about building more housing units, but also about the location of those units. The theory is not perfect, but the underlying concept is strong: People pay more for a better location, which drives up housing unaffordability in that location. The bid-rent curves in the model describe the competing market and why those with higher incomes outcompete the less advantaged. It becomes clear that disproportionate increase of unaffordable housing in urbanization is not just a result of low housing supply or low enough prices; rather it is a consequence of lack of housing in specific locations (Pickford, 2016).

The supply of low cost housing in the urban counties of California is extremely rare due to intense competition for space in these areas. If the competition for land increases, houses get less affordable; if houses get less affordable, those with higher incomes will prize out the lower income groups; if the lower income people are outcompeted for land it is likely they will not decide to move into these highly competitive areas. Florida (2017 p.22) refers to this as the 'urban land nexus'. However, it must be emphasized that also less urban and peripheral counties have experienced declining housing affordability rates over the last decades (California Association of Realtors, 2018a). Besides, housing prices in urban areas have been higher than other places for a long time: In 1950 they were already twice the national average. Remarkable, however, was that this quadrupled in 2000 (Gyourko, Mayer and Sinai, 2013). So, it can be assumed that especially urban counties are experiencing considerably fast rising housing prices. Throughout the years California has become extremely urbanized, and the higher paid jobs started to locate in the urban areas of the state. At this moment about fifty per cent of the state's population lives in four counties (Los Angeles, Orange, San Diego and San Bernardino), while an additional thirty per cent of the population lives in nine counties (Alameda, Santa Clara, San Mateo, San Francisco, Contra Costa, Sacramento, Ventura, Riverside and Fresno). For all of California's residents it was estimated that about ninety-four per cent live in urbanization, while just six per cent live in the rural areas. However, it must be acknowledged that rural is defined as "all territory, population, and housing units that are located outside of urban areas and urban clusters" (Ratcliffe et al., 2016 p.2). For California, only four counties can be defined as being entirely rural, while just another seven counties can be defined as predominantly rural. Thus, most of California's residents live in predominantly urban counties (University of California, 2017; Ratcliffe et al., 2016). To qualify as an urban area "the territory identified according to criteria must encompass at least 2,500 people, at least 1,500 of which reside outside institutional group quarters (U.S. Census Bureau, 2010). Every ten years the U.S. Census Bureau identifies urban areas based on new census data. All urban areas in California's counties are shown in figure 3.

Figure 3 Urban areas in California counties

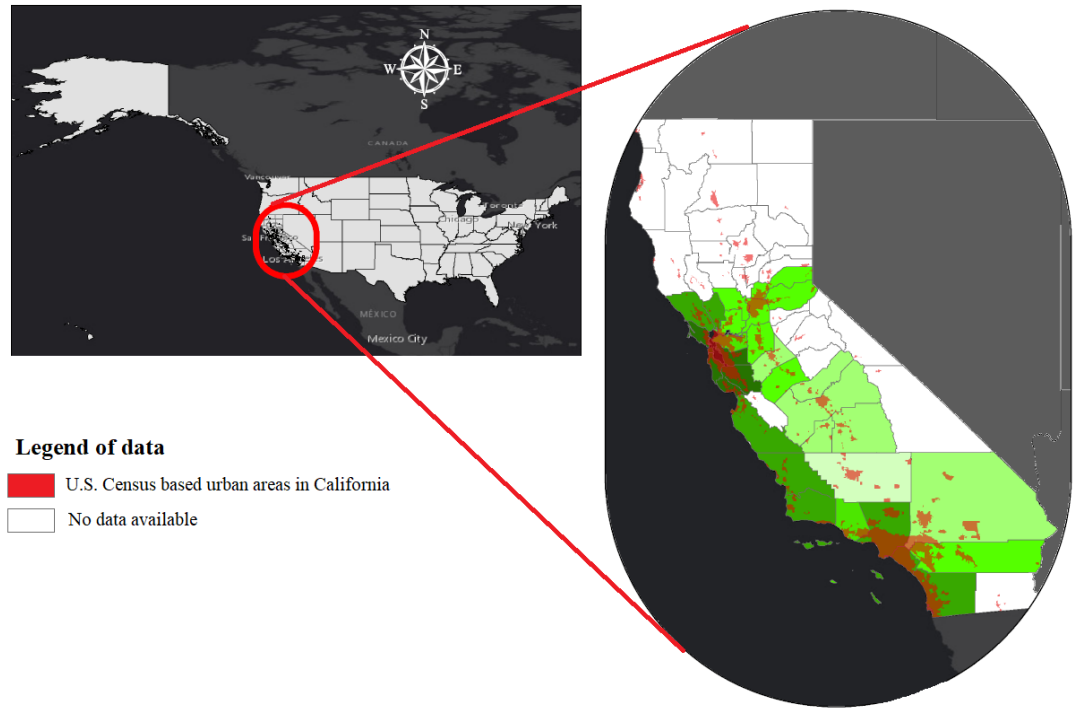


Source: Own design by the U.S. census urban/rural criteria (U.S. Census Bureau, 2010)

The downside of the development of the highly concentrated population in the urban counties and the fierce competition for urban land is that that housing prices are becoming disproportionally high and therefore unaffordable for some demographic groups. Employees in less innovative and successful industries, like blue-collar and service workers, are the most disadvantaged in this situation (Florida, 2017 p.6). But also growing numbers of economically advantaged people are being priced out from these expensive areas because they see their money disappear and fear that their children will never be able to afford these high costs of living. In her book about the rise of the new global super rich Freeman (2012, p.188-229) and supported by a 2016 London School of Economics study, it became clear that it are no longer just the less moneyed people who are being prized out from their houses and apartments; instead long-established elites and more affluent people in general are also losing out now. Super rich people – often wealthy foreign buyers – are buying houses from affluent individuals. However, who is going to feel guilty for the rich selling their homes to the super rich for huge sums and enormous profits? Most likely there won't be so many people feeling sorry for that. Yet, it highlights the fact that also more affluent people are facing the competition of an even higher order. Instead of starting a living in the houses they buy, the super-rich rather look for place to park their money. Next to the super-rich, there are also corporations, real estate investment trusts, hedge funds, and wealth funds investing huge amounts of their assets in these urban and desirable places (Florida 2017, p.39). In a study to the geography of the super-rich, Florida and Mellander (2016) found that the most desirable places for them to buy are located in the urban, high-tech places, which are the urban coastal counties of California. In 2016, California was the state with the third highest median monthly housing costs of the U.S., but it ranked number one considering the growth of these costs since 2012. The state also ranked second in the percentage of household income spend on housing. However, it is the urban counties that contribute mostly to the high housing costs, while the exurban and rural counties are reasonable affordable (Bean, Brown and Pullés, 2018; California Association of Realtors, 2018a; National Low Income Housing Coalition, 2017). How urbanization and housing affordability are related to each other becomes visible in figure 4, where a map was created of housing affordability for each of the California counties which were studied in the research of this paper. More information about the counties that were included in the study was provided in the third section of this paper. The figure contains the housing affordability rates of 2018 from the California Association of Realtors (2018a). The color ramp indicates the housing affordability in a county: The darker colors represent

the less affordable counties, while the lighter colors reflect the counties that are more affordable. It becomes gradually visible that the urbanized coastal counties are the ones that are least affordable, while the less urban counties that are located inland are more affordable.

Figure 4 Urban California and housing affordability in the state's counties



Source: Own design by the U.S. census urban/rural criteria (U.S. Census Bureau, 2010) and housing affordability rates of 2018 (California Association of Realtors, 2018a).

Note: The green color ramp represents the housing affordability in California counties (Darker= less affordable; Lighter = more affordable)

While the affluent people, high-tech industries and knowledge jobs moved to dense urban places, there is a risk that more affordable and less urban areas get more attractive for lower income groups (Florida, 2017 p.159-166). These less urban places are characterized by more square footage, less noise and pollution, and more privacy (Kurtz and Eicher 1958; Nelson and Sanchez 1997). In California, the urban counties are mostly located along the coastline whereas the less urban – or suburban and exurban – counties are based inland. While the urban counties attract people and jobs because of their convenience and strong productivity, farther-out locations in suburban counties offer space and affordable land. Because of that it is way cheaper and easier to seek residence in these less urban counties of California. Yet, the urban counties of California are the ones who create the high-tech, knowledge based jobs necessary for sustainable high-quality economic growth. Since it is not likely that the state's natural population will meet the demand on the labor markets, it is important that urban counties attract workforce talents from elsewhere. In his book about the new urban crisis, Florida (2017, p.166) cites that "too much of our precious national productive capacity and wealth is being squandered on building and maintaining suburban homes with three-car garages, and on the roads and sprawl that support them, rather than being invested in the knowledge, technology, and density that are required for sustainable economic growth. The suburbs aren't going away, but they are no longer the apotheosis of the American Dream and the engine of economic growth". Even though regional housing affordability could be an influential variable in explaining individual in-migration patterns in California, the internal migration decision can be influenced by many more factors. Therefore, there the other predictors that were used as control variables in the quantitative study of this paper were elaborated in the following subsection.

2.5 Other predictors

In order to make the quantitative analysis more robust and accurate, a series of control variables have been added to the regression models in the statistical analysis. A multinomial logit framework was used for the estimation of the influence of housing affordability on internal migration patterns into California counties. More information about the statistical methods was provided in section 3. There was chosen for c*ontrolled* multinomial logit models because it enables the researcher to control directly for confounding variables in the regression between housing affordability and internal migration. A controlled regression analysis basically implies that when looking at the effect of housing affordability on internal migration, also the effects of all other variables are taken into account. Controlling for confounding variables can be done by either including other predictors of the dependent variable in the regression or by making the other variables take on a fixed value (Long and Freese, 2006 p. 228; Mehmetoglu and Jakobsen, 2016 p.180). Both methods were applied in the statistical model of this research. As mentioned before, the individual's decision to migrate is a result from two components: The decision *to move* and the decision *where to move* (Roseman, 1977). For the first component, a wide range of predictors have been added to the multinomial logit models, which – according to available literature – also influence the individual decision *to move.* For the latter component, fixed values have been applied to control for any other regional factors that may influence the decision *where to move.*

To control for the decision *to move* a series of control variables have been added to the multinomial logit models. The ASEC data contains comprehensive cross-sectional data related to individual and household characteristics, socioeconomic status, and housing features. The controls that relate to individual and household characteristics are those that happen in the life-cycle (Roseman, 1977). Included were the year in which the individual internally migrated or not, age, gender, birthplace, marital status, number of children, and race as control variables in the analysis. Multiple studies have demonstrated that internal migration rates in the U.S. are procyclical – meaning that migration rates rise in good economic times and fall during economic downturns (e.g. Greenwood, Hunt and McDowell, 1986; Greenwood, 1997; Milne, 1993; Pissarides and Wadsworth, 1989). Therefore, the years between 2000 and 2018 have been grouped into three categories: The period 2000 to 2006 when GDP growth was relatively stable in California, the period 2007-2012 when the economy was in a financial recession and GDP growth fell substantially, and the period 2013-2018 when the state's economy recovered. In the U.S. younger adults have the highest propensities to move; the propensity to move decreases when age increases. The decision to migrate also tends to be lower for those with a racial background other than white, as it is for those who were born abroad, while individuals without children are more likely to migrate than those with one child or more. Meanwhile there are generally no differences between men and women and their internal migration propensities (Molloy, Smith and Wozniak, 2011), whereas single individuals are more likely to migrate than married couples (Clark and Lisowksi, 2017). Subsequently, the controls that relate to socioeconomic status can be seen as the events that happen in the individual's career cycle. The following socioeconomic variables have been included in the analysis: Education, household income, employment status and whether an individual worked last year or not. For the U.S., Molloy, Smith and Wozniak (2011) argued that the largest differences in the propensity to move are between individuals who are unemployed and those who are either employed or not in the labor force; between individuals who completed a college degree or higher and those with lower education levels. The last variable controlled for was tenure. This variable represents housing features and indicates whether the individual is a renter or homeowner. In her article about the two-sided relationship between population and housing, Mulder (2006) emphasizes the fact that home-ownership stands in the way of residential mobility and migration because it binds people to their place. For most countries, including the U.S., the likelihood of migration is much lower for homeowners than for renters (Speare, Goldsteind & Frey, 1995; Helderman, Mulder & Van Ham, 2006). This is partly because homeowners have to deal with much higher transaction costs related to moving compared to renters (Molloy, Smith and Wozniak, 2011).

To control for the decision *where to migrate* fixed county effects have been applied. It must be noted that there is only one variable that predicts the regional influence, namely the county's housing affordability, while everything related to the regional influence on internal migration was predicted by

this variable. Regional data related to internal migration in and between counties is rather scarce or not publicly accessible for the 2000-2018 periods. Therefore, there was chosen to add a fixed effect to the analysis by including all counties as categories in the multinomial logistic model (according to the method of Allison, 2009 p.28). This ensures that everything specific to a particular county and constant in time is captured in the analysis, such as the influence of quality of life within one county relative to another. Although using fixed effects is a rough method of modeling and despite the fact that it is not entirely clear what it exactly captures, it controls for every regional aspect that could also potentially influence internal migration. Remarkable is that the model does not change much after implementing the fixed effect, which suggests that nothing important is missing in the model. The standard errors for each county have been clustered to ensure that outcomes for people who live in the same county are not correlated. Even though the fixed effect was part of the multinomial logistic regression, there was chosen – in order to conserve space – to not include the county categories in the results in the fourth section.

2.6 Conceptual framework

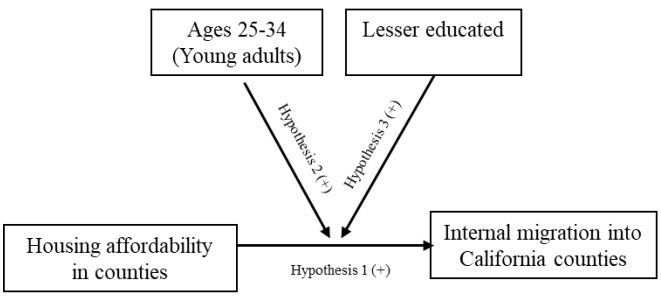
The characteristics of places have always played a major role in the internal migration choices of individuals in the U.S. (e.g. Lowry, 1966; Whisler et al., 2008). Regional characteristics and their influence on internal migration are of increasing interest among policy makers in California counties, including those that fear that internal migration is going to happen into more affordable counties (e.g. Orange County Business Council and Orange County Development Board 2017; 2018). Since internal migration has always been important for the (economic) development of California and because it is expected that internal migrants will be needed more than ever to guarantee a strong and sustainable future knowledge-based economy, understanding the role housing affordability has on the state's internal migration is useful to either facilitate or counteract on residential choice. This paper seeks to find empirical evidence regarding the influence of regional housing affordability on internal migration into the Golden State. Therefore, the main research question *'What is the influence of housing affordability on internal migration into California counties?'* was introduced in the introduction of this paper. In order to answer this question, three additional sub-questions were prepared.

The first sub-question relates to the extent in which housing affordability has influence on internal migration into California counties. Within California the housing affordability rates highly vary between its counties. Recent studies have found evidence that the state was already losing the lesser skilled for a while but also started to lose some of the higher skilled workers more recently, who move in a pattern that implies that housing affordability could be a driving force behind these moves (Bean, Brown and Pullés, 2018; Gunderson and Sorenson, 2010; Johnson, 2018). However, if housing affordability would be a driving force behind out-migration patterns from California, it would also be expected to have an influence on internal migration patterns into California. A multinomial logit framework was used for estimating the relative risks of internal migration by housing affordability on the county level (Mehmetoglu and Jakobsen, 2016 p.180). More information about the statistical methods was provided in section 3.3. Following recent evidence from literature it was expected that the more affordable counties in California have higher relative risks of internal migration into the county compared to the less affordable ones. Therefore, it was expected that the relative risks of migration compared to no migration into a county increase with housing affordability (hypothesis 1). The second sub-question asks how the relationship between housing affordability and internal migration is different between age groups. There is a growing interest in understanding the migration of the young adults (aged twenty-five to thirty-four), also referred to as the millennials (Benetsky and Fields, 2015; OCBC and OCDB, 2018). Losing this group of people may seriously harm and threaten future economic prosperity in some of California's counties, including those that rely on highly skilled and educated people. It is the young adults who have the highest migration propensities and are potentially those who react strongest on housing affordability differences in California. Although young adults might have high-paying tech jobs in the coastal urban California counties and are attracted more by urban amenities, they are also the ones who just start at the labor market, and are less likely to afford the costs of buying a decent house or renting a proper apartment in the expensive

counties of California. From economic location theory it can be argued that they generally do not have the resources to win the competition for urban land against older age cohorts with generally higher incomes and more assets. Moreover, evidence shows that internal migration propensities fall with age (Molloy, Smith and Wozniak, 2011; Roseman, 1977), which implies that older age groups are less likely to move than younger ones. Therefore, it was expected that relationship between housing affordability and internal migration into California counties is stronger for the young adults than for any other age group (hypothesis 2). When looking at educational background, it was expected that lesser educated people with generally lower incomes are not able to win the competition for expensive land against those who can afford to bid higher prices – that is, the higher educated. As a result, the lesser educated households would have to seek for residence elsewhere. If counties are not able to provide any affordable housing for lower income groups in specific locations, it is more likely that these people start looking for residence in other counties that are more affordable. If the budget constraint is limited, it is more likely that larger family households who are less educated and have lower incomes prefer a three-bedroom apartment in more affordable counties than a one bedroom apartment in a highly urbanized/competitive county. Besides, the less urban counties in California which surround the highly urbanized counties also began to offer many amenities. It is however the highly urbanized counties that have the strongest labor markets, and which are the most important for future economic prosperity (Florida, 2017 p.166). Since the higher educated households are likely to price out the lesser educated ones in California's coastal urban counties, the relationship between housing affordability and internal migration was expected to be stronger for lesser educated individuals than for those with a higher educational attainment (hypothesis 3).

By putting this all together, as California transitioned into one of the most innovative economies of the world, the most affluent residents started to move into the urban and coastal counties with the strongest high-tech labor markets. However, the less affluent ones – i.e. the younger and lesser educated individuals in this research - are falling further behind and cannot outbid the moneyed ones. Economic location theory explains how the lower-income groups are being outcompeted for urban land by the higher income groups. Expected is that the lower educated and younger individuals are the ones who are affected strongest by housing affordability rates in counties, and therefore strengthen the relationship between housing affordability and internal migration into California counties. For clarification, a conceptual model was prepared in figure 5. The conceptual model shows the expected relationships between the main variables of the study in this paper.

Figure 5 Conceptual model



Hypothesis 1: When housing affordability increases, the relative risks of internal migration - compared to no migration - into California counties increase

Hypothesis 2: The relationship between housing affordability and the relative risks of internal migration into California - compared to no migration - is stronger for young adults than for any other age group.

Hypothesis 3: The relationship between housing affordability and the relative risks of internal migration into California - compared to no migration - is stronger for the lesser educated than for higher educated people.

Source: Own design

What is the influence of housing affordability on internal migration into California? The remaining of this paper discusses, examines and elaborates on this question. Does housing affordability in counties influence the individual's decision where to migrate? Or is does housing affordability not influence internal migration into California at all due to – for example – the magnetic force of the strong high-tech labor markets in the coastal urban counties? The following section elaborates on the research design, while the empirical analysis and the results from the study in this paper were described in the fourth section.

Section 3 – Research design

After having provided the literature review in the previous section, this chapter continues to elaborate on the research methods used to answer the research questions prepared in the introduction of this paper. First, an explanation was given of the data that was used for the quantitative analysis in this research. Subsequently the statistical methods were discussed, while the last past in this section elaborates on the use of mixed methods in the study of this paper. The series of control variables used for the multinomial logistic models were already discussed in the previous section.

3.1 Data

Internal migration. Now the main objective of this paper was discussed in the previous sections, this subsection will elaborate on the data that was used for the study in this paper. However, data sources to explain migration patterns within the U.S. are rather scarce. There are only a few sources that provide micro data about migration variables, including the American Community Survey (ACS), the Current Population Survey (CPS) including the Annual Social and Economic Supplement (ASEC), the Internal Revenue Service (IRS), and the Federal Reserve Bank of New York's Consumer Credit Panel (CCP). The misery of migration data relates to problems of availability, quality, comparability and confidentiality. For internal migration data there are generally two persistent and basic problems: Limited geographic coverage and detail, and unrefined time intervals (Cooke, 2013; Isserman, Plane, and McMillen, 1982; Molloy, Smith and Wozniak, 2011). Concerning the limitations from geographic coverage both the ACS's and the CPS's & ASEC provide data at the county level, whereas data from the IRS and CCP are able to examine at smaller geographic units. Some researchers claim that migration operates simultaneously on multiple levels and is more often identified to just being related to the local level (Massay et al., 1999; Bilsborrow, 2016). They claim that smaller geographic units, such as cities, school districts and neighborhoods always should be included. Nevertheless, this research objects to estimate the influence of housing affordability in counties on internal migration into counties in California, and therefore this research does not suffer from the ACS's and CPS &ASEC's geographic limitations. As mentioned in the previous section, two important decisions have to be made to define migrants (Long, 1988 p. 4-12; Molloy, Smith and Wozniak, 2011):

1. The geographical unit of migration including the potential origin and destination locations
2. The time period in which individuals move between origins and destinations

In this research migrants are defined as those who either migrated from within or from another county into a California county, relative to where the same person lived in the previous year. That also means that a migrant could have moved from another state. According to Roseman (1977) it is critical to understand migration at the smallest geographic unit in order to thoroughly explain population change. County-level analyses are often used as a geographic unit of examination because it is often the smallest geographic unit for which data is readily available. The ACS migration data, however, suffers from its highly coarse time series because it only measures migration of an individual between the census date and 5 years prior to the census date. Since this research depends on annually changing housing affordability index rates, yearly migration data would be more valuable for this paper. Migration data from the IRS is based upon matching tax returns of tax filers in consecutive years. By computing the number of returns and the number of exemptions claimed, county-to-county migration flows can be determined. However, micro data from the IRS has not yet been made publicly available and cannot be purchased or accessed by outside researchers, which makes it impossible to use this data resource. In contrast to other data sources, the longitudinal character of the CCP data makes it possible to provide more nuanced observations of migration over the life course. The panel structure of the data combines both cross-sectional and time series data to determine how an individual or household has changed over time. However, the Reserve Bank of New York only allows Federal Reserve researchers

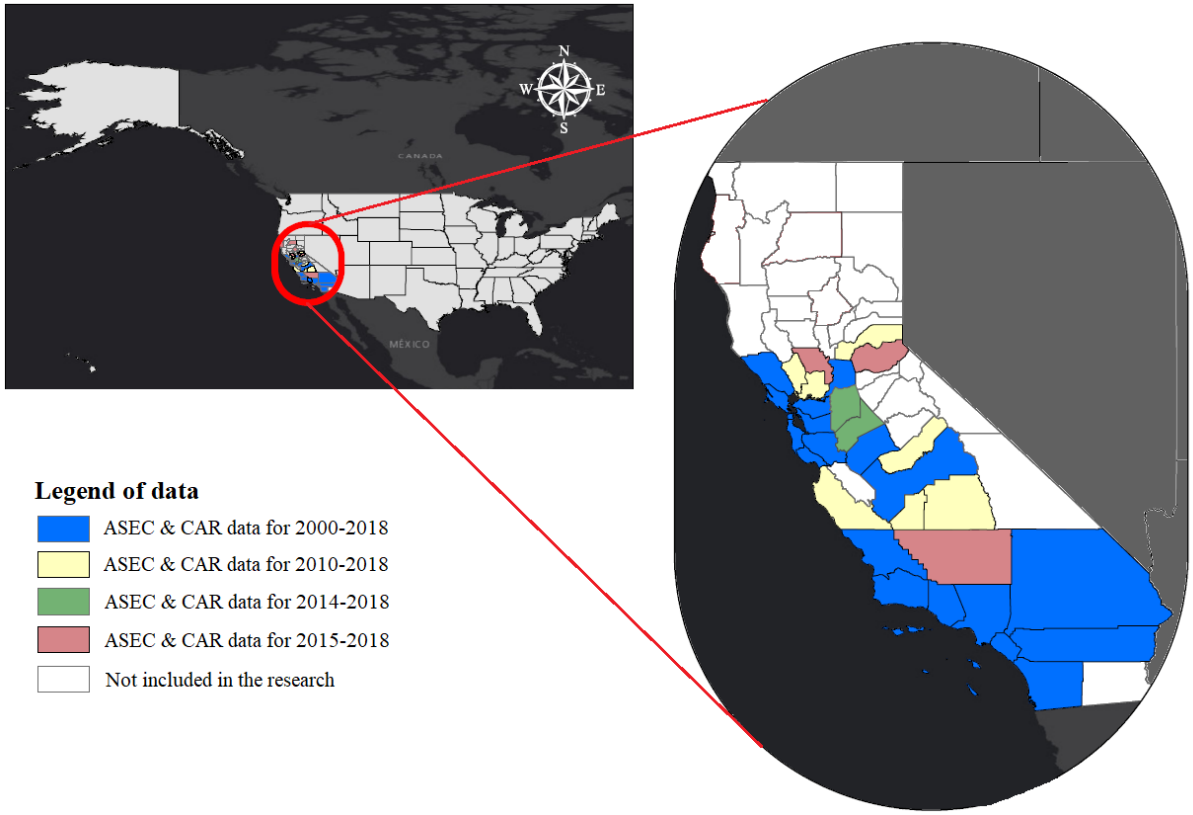
to access the CCP micro data due to the sensitivity of the data and the restrictiveness of their federal contract[4]. Therefore, it was also not possible to use the CCP data for the study in this paper. Nevertheless, after analyzing the different data sources from above, it became clear that the Annual Social and Economic Supplement (Flood et al., 2018) best fits the definition of migrants because using the ASEC data does not imply any geographical limitations concerning the study, and it has the most comprehensive cross-sectional data related to individual and household characteristics, socioeconomic status, and housing features. These are important control variables for the estimation of internal migration as explained in the previous section. Besides, the data does not have any limitations concerning unrefined time intervals because it is a yearly supplemental inquiry to the Current Population Survey (CPS) and adds special topics for the month March, every year. For this research the yearly ASEC releases of 2000 until the most recent release of September 2018 have been used to perform further analyses. To make cross-time comparisons using the data, variables were coded identically for all years. Flood et al. (2018) emphasize that the data is not a collection of compiled statistics, but instead it only contains microdata. Individuals in de database have been organized into households, which enables studying the characteristics of people in the context of their families or other co-residents. Due to confidentiality reasons the data is only limited to a respondent's current county of residence, which is the destination county after changing residence. The origin county of the migrants cannot be observed, which means that this research is limited to the examination of destination counties. For identifying a respondent's origin county of residence access to restricted use data is necessary, which is only available through US Census Bureau Research Data Centers. These are designated locations with a standard set of security measures in place that ensure the confidentiality of restricted use data within the center. There is an extensive application process to be approved to use these data and, in many cases, the work must be done inside one of the Research Data Centers. However, after applying for admission to the data center, it became clear they did not grant access to outside researchers. Nevertheless, only studying the destination county of internal migrants in the ASEC does not limit this research's aim to estimate internal migration into a county by regional housing affordability. Therefore, it can be stated that the data limitations of the ASEC do not restrict the analyses in the study of this paper. In their research about internal migration of the college educated and quality of life indicators, Whisler et al. (2008) argued that there are two possible methods to quantify internal migration patterns: Either by determining the choices made by individual movers or to assess aggregate movements of people. In this paper there has been built on the first method of determining the migration choice by individual movers in the ASEC data. The migration variable, the main predictor variable in the analysis, has been modified by using geographic data concerning individuals and households in the ASEC, and migration variables from the survey that identify whether respondents moved and if so, if they moved from within a county, between counties, between states, or abroad . Those who were living in the same house as one year ago were considered non-movers and were not asked any more questions related to migration over the past year. The modified migration variable indicates whether the respondent was living in the same house; moved in from within the same county in California; or moved in from another county relative to the previous year.

Housing affordability. The main predictor variable to estimate the individual's decision to internally migrate into California counties is housing affordability. To measure housing affordability in counties in California, index rates from the CAR, the California Association of Realtors (2018a), were used. The CAR computes a monthly or quarterly Traditional Housing Affordability Index for most counties in California, which measures the percentage of households that can afford to purchase a median priced home in California counties based on traditional assumptions (California Association of Realtors, 2018b). Until 2006 the housing affordability index rates have been calculated for every month, while from 2006 the rates where only computed quarterly. At the time of performing the statistical analysis the housing affordability index rates for the last quarter of 2018 were not yet available. The index rate has been averaged for every year for every county represented in both the ASEC and the CAR. Important is to notify that not every county was represented in the ASEC, and that the CAR did not compile a housing affordability index rate for every county for a

[4] This information became clear after having email contact with the Reserve Bank of New York

particular year because they did not yet have data from those counties back then. Even though the association started to include the smaller counties in their tables from 2010, not yet all counties are represented for the years between 2000 and 2018. Furthermore, the research was limited to those counties in relative close proximity to urbanization. Since California has always symbolized opportunity, and internal migration into the state was often combined with supplying the labour market, it would not make sense that migrants would move to the rural counties in the north for housing affordability reasons. These counties are far from strong labour markets and urbanization, and therefore have way less amenities than those at a closer proximity to the urban counties. For that reason Butte, Humboldt and Shasta Country were also excluded from the analysis. For clarification, the counties and years that were included in the analysis are shown in figure 6. A table of the counties that were included in the statistical analysis can be found in Appendix A. Another table with the housing affordability rates per county was provided in Appendix B, while appendix C describes the methodology used by the CAR to compute their monthly/quarterly housing affordability index for each county (California Association of Realtors, 2018b). There are some limitations concerning the method by which the Housing Affordability Index has been calculated. The methodology is based upon assumptions that do not necessarily have to apply for all respondents in the ASEC. Besides, the index is established by the C.A.R.'s monthly existing home sales survey, and therefore does not include all homes sold. Nevertheless, the survey is based on reports of closed escrow sales from more than 80 boards and multiple listing services around the state, consisting of a large share of the home purchases in California. Despite its limitations, the Traditional Housing Affordability Index from the CAR is one of the best fundamental measures of housing well-being for buyers in the state, and therefore provides a great deal in representing the county's housing affordability. To analyze the influence of these housing affordability index rates on internal migration into California counties, multinomial logistic regression models – discussed in section 3.2 and elaborated in the fourth section– were performed by using the statistical package Stata.

Figure 6 Available data from California's counties



Source: Own design by data from the Annual Social and Economic Supplement (Flood et al., 2018) and California Association of Realtors (2018a)

3.2 Statistical methods

As mentioned before, there are two ways to analyze internal migration patterns, either by 1) determining the choices made by individual movers or 2) assessing aggregate movements of people. In this study there is chosen to determine individual choices of ASEC respondents and estimate the influence of a county's housing affordability on internal migration patterns. Since the dependent variable in this study contains three independent categories – 1) no internal migration, 2) migration within the county, and 3) migration to another county – a multinomial logit framework was used for estimating the risks of internal migration by housing affordability on the county level. In their book about applied statistics for the social sciences, Mehmetoglu and Jakobsen (2016, p.180) describe that multinomial regression builds on ordinary logistic regression but provides the researcher with the possibility to study dependent variables that have more than two not naturally ordered categories. The key assumption for multinomial logit frameworks is that the observations must hold the *independence of irrelevant alternatives assumption* (also known by the *IIA property assumption*) which implies that observations and its errors are independent from irrelevant alternatives. Basically, the assumption holds that the dependent categorical variable – internal migration in this study – must be mutually exclusive and exhaustive by including all possible alternatives as values. However, there is no good applied test for this assumption. According to Long and Freese (2006, p.407) the existing tests – consisting of the Hausman test and the Small-Hsiao test – show inconsistent results and dubiously judged simulations. Therefore, they claim that the researcher should be able to plausibly assume that the alternatives are distinct and weighted independently. Since the internal migration variable contains three fully independent and exhaustive alternatives, it can be assumed that the multinomial logit framework in this study holds the IIA assumption.

The multinomial logit models compare the risks of internal migration into a county from within the same county or from another county relative to no internal migration. Nine models have been estimated in order to understand the relationship between housing affordability and internal migration among different age groups. The first model includes the housing affordability index and the individual and household predictors. In the second model the socioeconomic and housing predictors have been added, to determine whether these controls improve the model as a whole. The first two models are general models that apply for all ages. Models three to six are specified for specific ages in order to estimate whether there was a difference between the risks of internal migration by a county's housing affordability between different age groups. The third model only includes the 15-19 and 20-24 age groups, the fourth model focuses on the 25-29 and 30-34 age groups, the fifth model takes the 35-39 and 40-44 age groups into account, and the sixth model is specified for all other ages. The last three models were specified for different education levels. The seventh and ninth models were specified for respectively the lesser and higher educated. To define lesser and higher education, the paper of Bean, Brown and Pullés (2018) was followed where the lesser educated are those with a high school degree or less, while the higher educated have at least a four year college degree – or bachelor's degree. In addition, model eight focused on the middle educated, whose services might be of great importance to sustain California's future economy. The middle-skilled jobs require more education and training than high school but less than a four year college degree - or bachelor's degree (Fuller and Raman, 2016). In all models the 'no migration' category was used as reference group, comparing their housing affordability index rate to that of internal migration from the same county or from another county. From the equations below it becomes clear that different logit coefficients can be obtained for two of the three categories of the dependent variable. The first equation (3.1) computes the relative risk of internal migration from the same county in California (P_w) compared to no migration (P_n). The second equation (3.2) calculates the relative risk of internal migration from another county or another state (P_o) relative to no migration (P_n). For both equations the relative risks are calculated by the coefficients of the independent variable, the coefficient of the other predictor variables and the standard errors. Eventually, the combined probabilities of P_n, P_w and P_o equal 1 (3.3). The basic equations were used for the estimations in section 4 of this paper. It should be noted that the outcomes of the multinomial logit models are relative risk ratios, but can be interpreted as odds ratios even though they are not exactly the same. However, odds, or odds ratios, are usually being interpreted as being equivalent to the relative risk ratios. In their assessment about odds ratios, Davies, Crombie and

Tavakoli (1998) explain that if the odds ratio is interpreted as a relative risk, it will always exaggerate any size of effect. Nevertheless, serious distinction between the odds ratio and relative risk ratio only occurs with particularly large effects on groups that have a high initial risk. Therefore, interpreting the odds ratio as a relative risk is usually unlikely to be seriously in error. Since the results in section 4 do not show extremely large effects regarding a one rate increase of the housing affordability index on internal migration compared to no migration, interpreting the odds as relative risks is unlikely to lead to serious overstatements in this study's estimations.

$$\frac{P_w}{P_n} = b_{0w} + b_{1w}x_1 + \ \ldots + b_{iw}x_i + e_w \qquad (3.1)$$

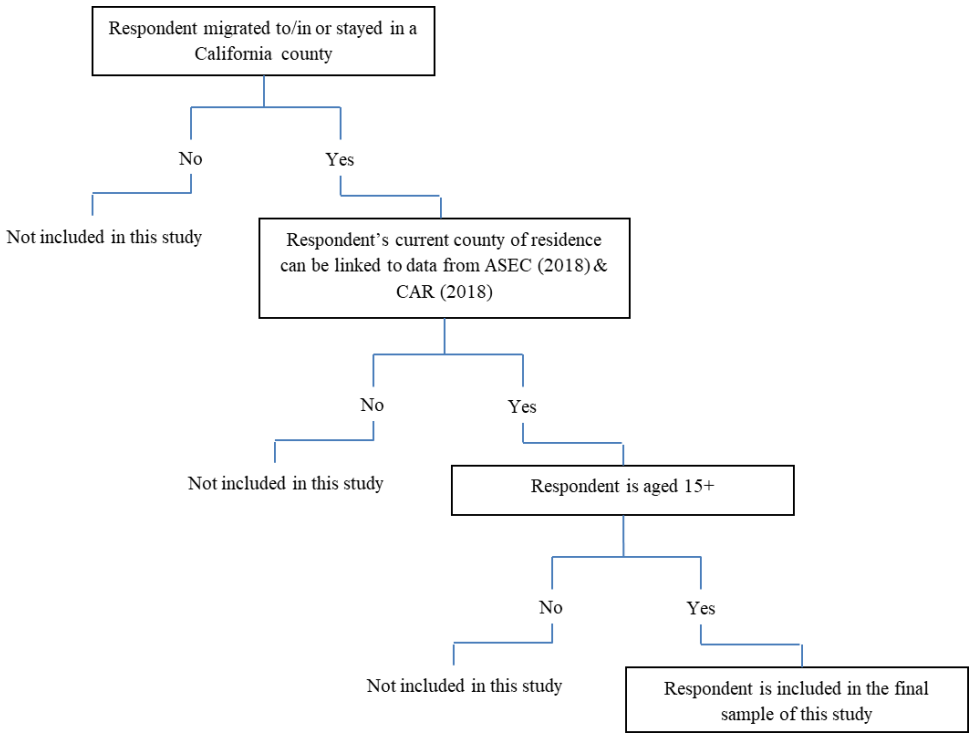
$$\frac{P_o}{P_n} = b_{0o} + b_{1o}x_1 + \ \ldots + b_{io}x_i + e_o \qquad (3.2)$$

$$P_n + P_w + P_o = 1 \qquad (3.3)$$

Source: Own elaboration on Mehmetoglu and Jakobsen (2016, p.181)

The data contains observations of 3,807,919 different individuals for eighteen different years. However, the modeling universe limits the observations since the research narrows down to those currently living California's counties, whose residential status has either changed or not compared to the year before. That means those who moved abroad were not included in the modeling universe. As mentioned in section 3.1, Butte, Humboldt and Shasta County have also been removed from the analysis, which reduces the modeling universe with 751 observations. Excluding these rural counties from the analysis does not influence the coefficients and their significance. Besides, these counties did not have many observations. Therefore, it can be stated that limiting the analysis to the other counties did not influence the results. In addition to these limitations and those related to the housing affordability data as mentioned in the previous section, the universe is further limited to individuals that fit the description of the other predictor variables. The ASEC only includes persons with the age of 15 years and older in the variables *employment status* and *education*, since younger respondents do not meet the eligibility requirements for these variables. For example, respondents under the age of 15 are not likely to have entered the labour force yet. Collectively, the data-based limitations reduce the modeling universe in the full model to 213,503 observations, as shown in table 2 in section 5. The sample selection is shown in figure 7. Conform other internal migration papers, a 95% confidence interval level was used to test for significance between the different variables (among other Clark and Lisowski, 2017; Saks and Wozniak, 2011; Whisler, 2008). The syntax of the statistical analysis was attached in Appendix D. In the next section the results of the estimation from the logit frameworks will be discussed, followed up by the conclusion and discussion in section 5.

Figure 7 Sample selection



Source: Own design

3.3 Mixed methods: Adding qualitative aspects to quantitative data

Even though this research mostly relies on quantitative data, the last sub-question in this paper relates to the housing affordability mechanism and how it works on internal migration in California. It must be emphasized that the core of this paper does not rely on obtaining a qualitative understanding about housing affordability and internal migration in California; rather it is to estimate migration from housing affordability in its counties. Therefore, it can be stated that this paper mainly relies on quantitative data. However, the importance of gaining qualitative knowledge about this topic cannot be neglected. The first three sub-questions apply to the estimation of housing affordability on internal migration into the state, and thus relate to quantitative data. So far it does not provide much information about the larger mechanism behind housing affordability in California and how it drives internal migration. Therefore, there was chosen to apply a qualitative aspect that related to the last sub-question in this research. This should help to understand the larger framework in which housing affordability could or could not drive internal migration in California. Combining quantitative- and qualitative methods is known by the methodology of mixed methods (Winchester, 1999), which can be a proper technique in order to get a broader understanding of mechanisms such as housing affordability. Applying qualitative aspects to a quantitative research can help to assist in tackling the weaknesses of the latter, by using the strengths of the first (Sale, Lohfeld and Brazil, 2002). According to Rabionet (2011) qualitative methods are a useful and flexible tool to capture real experiences from people, instead of only looking at quantitative data. Quantitative research, however, is valuable for obtaining information about the relationship between different variables (Carr, 1994). Since evidence shows that qualitative research can add sufficient value to information obtained from quantitative data, there was chosen to combine both techniques known as the approach of mixed methods.

The qualitative information for this paper was obtained from two semi-structured interviews, which is an instrument to explore the quantitative data. Longhurst (2010, p.103) describes a semi-structured interview as a "verbal interchange where one person, the interviewer, attempts to elicit information

from another person by asking questions". Even though semi-constructed interviews are usually conducted with a predetermined list of questions, these unfold in a conversational manner offering participants the possibility to explore issues they feel are important. Because of that, semi-structured interviews can be a valuable tool to obtain 'hidden', or undiscovered, knowledge. According to Longhurst (2010) this kind of interviews are used to collect data on a diverse range of subjects. It is useful to examine complex behavior, opinions, emotions and experiences. The way how participants are selected for semi-structured interviews is of vital importance. According to Cameron (2005) participants are usually chosen on the basis of their experience concerning the research topic. Choosing a random, representative and objective sample is strictly important for quantitative analyses, but this is not the case for qualitative research methods. The aim is not to be representative, but to understand how participants experience issues related to the topic (Valentine, 2005 p.111). For the research in this paper there was chosen to conduct two interviews with researchers who had different research interests: Academics versus business. This was done in order to get a more comprehensive understanding of housing affordability in California counties and how it works on internal migration into the county. The first interview was conducted with Prof. Susan K. Brown, professor of Sociology at the University of California in Irvine. Her research expertise lies in (international) migration, educational inequality and urban sociology. Among other she examines how people migrate into the U.S. and California across multiple dimensions, such as education, income, and residential location. Recently she wrote an article about migration in California together with F. Bean and S. Pullés "*Migration and the California Dream: Past, present and future"* (Bean, Brown and Pullés, 2018). In the second interview, Dr. Wallace Walrod was asked questions about housing affordability and internal migration in California, who has been in charge of research at the Orange County Business Council for more than twenty years. He was important for different projects with organizations such as Toshiba America Information Systems, John Wayne Airport, and Orange County Transportation Authority. Besides, he is founding president of the Center for a New California, which has a major role in shaping the state's future economic future by focusing on eminent issues including economic development, workforce development and infrastructure. An interview guide was provided in Appendix E. The results related to both the quantitative- and the qualitative part were provided and discussed in the next section. Nine different regression models were produced including their accompanying figures to answer the sub-questions related to the estimation of housing affordability on internal migration into California. Additionally, a summary of the relevant information from the interviews was provided to support the last sub-question relying on qualitative information.

Section 4 – Empirical analysis and results

This section elaborates on the previous sections of this paper by providing an empirical analysis, and by describing and explaining the results from the multinomial logistic regression models and the semi-constructed interviews. The syntax used for the quantitative analysis with the statistical package Stata was provided in Appendix D.

4.1 Descriptive results

The data in table 2 shows that the distribution of the main- and other predictor variables varied across the modeling universe. From the distribution of the modeling universe the relative share of respondents that did not migrate, migrated from the same county or migrated from another county have been calculated for each of the categories of the categorical predictor variables. This enables us to say something about the relative share of respondents within a category of a particular predictor variable and their migration status. As mentioned in the previous section, the modeling universe contains 213,503 respondents. Specific emphasis was placed on estimating the influence of housing affordability on internal migration patterns of different age- and education groups, which will be elaborated more extensively in the next two subsections. When looking at age in this study, it is the younger age groups who have a relative high share of internal migrants, whereas the migration portion seems to decrease when people get older. This is in line with the literature review in section 2, where internal migration propensities fall with age (Molloy, Smith and Wozniak, 2011; Roseman, 1977). The

percentages of movers for each age group were plotted in figure 8. Notably, there is a clear increase in the graphs of both internal migration from the same county and internal migration from another county for the younger age groups. Thus, as expected there seems to be a relation between internal migration and the younger age groups. Internal migration propensities do not seem to be very different between the various levels of educational attainment as the percentages of internal migrants are more or less similar. Yet, this does not tell anything about the influence of housing affordability on the internal migration patterns for these groups. When looking at the descriptive results of housing affordability, there does not seem to be much difference between the housing affordability rates and internal migration status. However, the descriptive statistics do not provide much information about housing affordability. To estimate the influence of housing affordability on internal migration within California, multinomial logistic regressions models have been prepared in the next two subsections. For the predictor variables related to individual and household characteristics, native born Americans, those who are alone in their marital status and those that do not have children also seem more likely to internally migrate. For the socioeconomic predictors, internal migration seems to be more popular among the higher educated and respondents with lower incomes. The share of respondents that reported they were unemployed, seem to migrate most often compared to other employment statuses, whereas those who did not work last year have higher shares of internal migration compared to those who had a job. And again, as expected from previous literature, home owners are less likely to migrate than are renters. Now the descriptive results were illustrated it is time to move on to the estimation models of the individual's decision to internally migrate into California counties by housing affordability in the next two subsections.

Table 1 Descriptive statistics for dependent, independent and control variables in the analysis.

	Modeling Universe		No migration		Internal migration from same county		Internal migration from another county	
Total	213,503		192,674		15,476		5,353	
Continuous variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Housing Affordability Index	31.4	15.1	31.4	15.1	31.5	15.4	31.0	16.2
Categorical variables – Individual and household characteristics	*N*		*N*	*Per cent*	*N*	*Per cent*	*N*	*Per cent*
Year								
2000-2006	68,323		60,502	88.55	5,529	8.09	2,292	3.35
2007-2012	72,700		65,532	90.14	5,528	7.60	1,640	2.26
2013-2018	72,480		66,640	91.94	4,419	6.10	1,421	1.96
Age group								
15-19	22,319		20,231	90.64	1,570	7.03	518	2.32
20-24	19,153		15,528	81.07	2,732	14.26	893	4.66
25-29	18,979		15,424	81.27	2,657	14.00	898	4.73
30-34	20,448		17,635	86.24	2,093	10.24	720	3.52
35-39	21,391		19,124	89.40	1,698	7.94	569	2.66
40-44	21,097		19,353	91.73	1,316	6.24	428	2.03
45-49	20,036		18,619	92.93	1,093	5.46	324	1.62
50-54	17,743		16,680	94.01	782	4.41	281	1.58
55-59	14,426		13,656	94.66	540	3.74	230	1.59
60-64	11,199		10,682	95.38	337	3.01	180	1.61
65-69	8,679		8,300	95.63	262	3.02	117	1.35

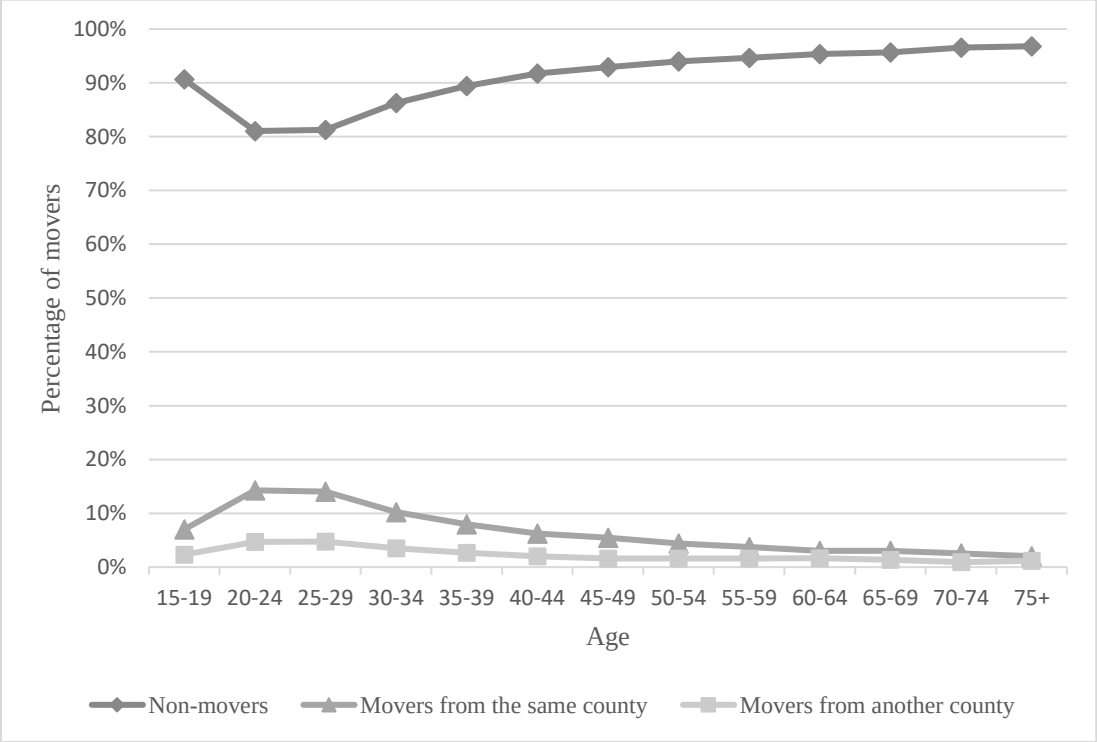
70-74	6,400	6,179	96.55	162	2.53	59	0.92
75+	11,633	11,263	96.82	234	2.01	136	1.17
Gender							
Male	103,056	92,824	90.07	7,556	7.33	2,676	2.60
Female	110,447	99,850	90.41	7,920	7.17	2,677	2.42
Birthplace							
Native born	130,334	116,847	89.65	9,783	7.51	3,704	2.84
Other	83,169	75,827	91.17	5,693	6.85	1,649	1.98
Marital status							
Married, spouse present	105,288	97,415	92.52	5,869	5.57	2,004	1.90
Married, spouse absent	3,858	3,385	87.74	322	8.35	151	3.91
Separated	5,271	4,526	85.87	581	11.02	164	3.11
Divorced	16,751	14,864	88.74	1,367	8.16	520	3.10
Widowed	9,497	9,010	94.87	346	3.64	141	1.48
Never married/single	72,838	63,474	87.14	6,991	9.60	2,373	3.26
Children							
None	117,178	104,453	89.14	9,189	7.84	3,536	3.02
At least one	96,325	88,221	91.59	6,287	6.53	1,817	1.89
Race							
White	162,483	146,910	90.42	11,759	7.24	3,814	2.35
Black/negro	13,621	12,066	88.58	1,155	8.48	400	2.94
Other	37,399	33,698	90.10	2,562	6.85	1,139	3.05
Categorical variables – Socioeconomic and housing characteristics							
Education							
Lesser educated	127,811	115,729	90.55	9,224	7.22	2,858	2.24
Middle educated	70,639	62,972	89.15	5,483	7.76	2,184	3.09
Higher educated	65,951	59,193	89.75	4,472	6.78	2,286	3.47
Household income							
Lowest quintile	67,961	60,276	88.69	5,910	8.70	1,775	2.61
Central quintile	67,250	60,524	90.00	4,960	7.38	1,766	2.63
Highest quintile	78,292	71,874	91.80	4,606	5.88	1,812	2.31
Employment status							
Unemployed	9,870	8,528	86.40	952	9.65	390	3.95
Employed	126,426	112,954	89.34	10,153	8.03	3,319	2.63
Not in the labor force	77,207	71,192	92.21	4,371	5.66	1,644	2.13
Worked last year							
Yes	137,106	122,179	89.11	11,110	8.10	3,817	2.78
No	76,397	70,495	92.27	4,366	5.71	1,536	2.01
Tenure							
Owner	122,754	116,911	95.24	4,246	3.46	1,597	1.30

Renter	90,749	75,763	83.49	11,230	12.37	3,756	4.14	

Source: Survey data from the Annual Social and Economic Supplement (Flood et al., 2018) and California Association of Realtors (2018a)

Note: M indicates mean, SD indicates standard deviation, and N indicates the number of observations

Figure 8 Internal migration by age group



Source: Own elaboration on survey data from the Annual Social and Economic Supplement (Flood et al., 2018)

4.2 Estimated internal migration differences by housing affordability from multinomial logistic regression models [General and age-specified models]

The results from the multinomial logistic regressions of the first six models can be found in table 3, while the results of the other models were provided in table 4. The results in both tables are represented by the adjusted relative risk ratios of internal migration from the same county or internal migration from another county relative to no migration (=reference group) by housing affordability and all other predictors controlled for in the model. As explained in the previous section, nine models were prepared in order to estimate the influence of housing affordability on internal migration. The independent variable and the individual and household predictors were included in the first model, while the remaining predictors were added in the second model, which shows an improved fit of the second- over the first model (Pseudo R^2 increases from 0.0519 to 0.1010). In both models the relative risk ratios of internal migration relative to no migration increased significantly when the housing affordability of the destination county increased. In both models the individual's relative risks of internal migration from the same county compared with no migration increased with respectively 1.005 and 1.004 (at a correspondently 99.9% and 99% confidence interval) when housing affordability increased with one per cent in the county, while the individual's relative risks of internal migration from other counties (or other states) into the county versus no migration increased with respectively

1.007 and 1.005 (both at a 99.9% confidence interval) with a one per cent increased housing affordability in the destination county (consistent with Hypothesis 1).

The first two models already provide an interesting indication that the younger age groups are indeed more likely to internally migrate into California counties than older ages, as expected from the literature review in the second section of this paper and the descriptive results from the previous subsection. The age group of twenty-five to twenty-nine years old (reference, and referred to as young adults) had significantly higher risks of internal migration compared to no migration relative to the other age groups in model 1 and model 2. Only in the second model was the relative risk ratio of the twenty to twenty-four years old age group significantly higher than the twenty-five to twenty-nine years old age group for internal migration from another county relative to no migration at a 95% confidence interval. Even though the coefficients of the thirty to thirty-four years old age group (also seen as young adults) were lower than the twenty-five to twenty-nine year old age group, they had significantly higher relative risk ratios than the older age groups compared to the reference age group for internal migration versus no migration. Yet, it is not possible to say anything about the influence of housing affordability on internal migration for these age groups. Therefore, models 3, 4, 5 and 6 were clustered for specific age groups in order to estimate the influence of age on the relationship between housing affordability and internal migration. Model 3 – which focused on the ages fifteen to twenty-four – shows that the influence of housing affordability was only significantly different from zero for internal migration from the same county compared to no migration. Remarkable was that when housing affordability increased with one per cent, the relative risk ratio of internal migration from the same county decreased with 0.996 relative to no migration for these age groups (at a 99% confidence interval). This indicates that individuals from this age group are less likely to migrate within the same county when houses get more affordable. Model 4 focused on the young adults (aged twenty-five to thirty-four), and finds significantly increased risks for both internal migration from the same county (1.009 at a 99.9% confidence interval) and internal migration from another county (1.007 at a 99% confidence interval) compared to no migration when housing affordability increased with one per cent. For the ages thirty-five to forty-four the relative risks of internal migration – both from within the same county and from another county compared to no migration – were not significantly different from zero when the housing affordability increased. Therefore, housing affordability does not significantly influence their risks of internally migrating versus not migrating into California counties. In the final age-specified model a one per cent increase in housing affordability only significantly increased the relative risk of migration from the same county (1.007 at a 99% confidence interval) relative to no migration for the ages forty-five and over. The influence of housing affordability on internal migration from other counties versus no migration was not significantly different from zero. When comparing the age-specified models (model 3, 4, 5 and 6) , the relative risks of internal migration into the same California county compared to no migration by a one per cent increase in counties' housing affordability increased most for the twenty-five to thirty-four years old age group, while this age group also had the highest relative risk of internal migration from another county (or another state) into a California county relative to no migration when houses in the destination county got more affordable (consistent with hypothesis 2). The fundamental assumption related to the above mentioned individual internal migration results by housing affordability implies that the other variables in the model are held constant.

Each multinomial logistic model shows positive Wald Chi^2 scores implying that the predictors in the model have improvement of fit over a model with no predictors (Mehmetoglu and Jakobsen (2016 p.173). The Pseudo R^2 measures to what extent the predictors in the model are able to explain the variance of the dependent variable. In most migration models the R^2 has a value between 0.1 and 0.2 (e.g. Clark and Lisowski, 2017; Saks and Wozniak, 2011). For this research only the second and third model have R^2 values between 0.1 and 0.2. However, according to Mehmetoglu and Jakobsen (2016, p.175) one should be careful with judging the goodness of fit by the value of pseudo R^2 in logistic models because in these models the value of R^2 is usually very low compared to its value in linear regression. When comparing the first model with the second model, the R^2 of the second model is almost double as high as the first. This indicates that adding the socioeconomic- and housing control variables to the individual- and household controls contributes to the comprehensiveness of the model as a whole by explaining more variance of the dependent variable. Table 3 provides the results of the

multinomial logistic regression of the general models (model 1 and 2) and the models specified for different age groups (model 3, 4, 5 and 6). The estimated relative risk ratios/odds ratios of internal migration compared to no migration by housing affordability – for the first six models – are shown in figure 9. The figure only shows the significant results. Our first sub-question asked to what extent housing affordability in California counties would influence the decision to migrate into these counties for the whole population. Models 1 and 2 were produced in order to answer this question; these show significant increased relative risks of internal migration (from the same and from other counties) versus no migration when housing affordability increased in the destination county. This confirms our first hypothesis arguing that when houses get more affordable, migration propensities into or within California counties increase. However, it is unlikely that the migration choice of the whole population is affected by housing affordability increases in counties. Therefore, the models were also specified for specific age groups. These models indicated that the young adults (aged twenty-five to forty-four) had the highest relative risks of internal migration into and within California counties relative to no migration when destination counties got more affordable housing rates. This is consistent with Hypothesis 2 related to the second sub-question, where it was expected that the relationship between housing affordability and internal migration into or within a California county would be stronger for the young adults than for any other age group. Even though there are some real significant results in this model, the relative risk ratios look rather small at a first glance. This is indeed the case when the housing affordability in the destination county only increases with one per cent. However, when the housing affordability of the destination county increases with *x* per cent, the relative risk ratio increases with $\exp(\beta)^{*} x$ – where $\exp(\beta)$ is the relative risk ratio estimated in the multinomial logistic regression (table 3). For example, the difference between Riverside County and Orange County was 18% in 2018 (California Association of Realtors, 2018a), which would imply that the relative risk of internal migration into Riverside County from Orange county already would increase with 1.124 for a young adult. Figure 9 illustrates this in a graph, where the relative risk ratios increase by housing affordability rates.

Table 2 Estimated Relative Risk Ratios from multinomial logistic regression for internal migration in California counties.

	Model 1 (N=275,656)		Model 2 (N= 213,503)		Model 3 (N= 41,472)		Model 4 (N= 39,427)		Model 5 (N= 42,488)		Model 6 (N= 90,116)	
	All ages		All ages		Age 15-24		Age 25-34		Age 35-44		Age 45+	
Predictor	Internal migration from same county	Internal migration from other county	Internal migration from same county	Internal migration from other county	Internal migration from same county	Internal migration from other county	Internal migration from same county	Internal migration from other county	Internal migration from same county	Internal migration from other county	Internal migration from same county	Internal migration from other county
Housing Affordability Index												
	1.00574***	1.00738***	1.00423**	1.00514***	0.99647**	1.00229	1.00878***	1.00689**	1.00523	1.00481	1.00731***	1.00811
	(0.00211)	(0.00185)	(0.00184)	(0.00196)	(0.00164)	(0.00287)	(0.00256)	(0.00285)	(0.00321)	(0.00339)	(0.00242)	(0.00651)
Year (ref: 2000-2006)												
2007-2012	0.856***	0.523***	0.830***	0.512***	0.931	0.554***	0.805***	0.602***	0.759***	0.442***	0.828*	0.417***
	(0.0476)	-0.036	(0.0485)	(0.0360)	(0.0437)	(0.0454)	(0.0385)	(0.0543)	(0.0729)	(0.0385)	(0.0840)	(0.0766)
2013-2018	0.710***	0.492***	0.662***	0.440***	0.586***	0.352***	0.665***	0.493***	0.701***	0.410***	0.708***	0.471***
	(0.0353)	-0.039	(0.0288)	(0.0317)	(0.0507)	(0.0448)	(0.0292)	(0.0483)	(0.0346)	(0.0591)	(0.0192)	(0.0533)
Age group (ref: 25-29)												
15-19	0.430***	0.382***	0.635***	0.645***								
	(0.0161)	(0.0303)	(0.0262)	(0.0636)								
20-24	0.984	0.915*	1.108	1.094*								
	(0.0729)	(0.0459)	(0.0698)	(0.0530)								
30-34	0.696***	0.740***	0.720***	0.758***								
	(0.0172)	(0.0357)	(0.0194)	(0.0402)								
35-39	0.518***	0.549***	0.581***	0.614***								
	(0.0154)	(0.0284)	(0.0291)	(0.0381)								
40-44	0.389***	0.400***	0.474***	0.486***								
	(0.0139)	(0.0258)	(0.0226)	(0.0298)								

45-49	0.329***	0.302***	0.431***	0.397***								
	(0.0183)	(0.0279)	(0.0341)	(0.0356)								
50-54	0.257***	0.274***	0.351***	0.379***								
	(0.0174)	(0.0269)	(0.0288)	(0.0398)								
55-59	0.210***	0.257***	0.305***	0.381***								
	(0.0107)	(0.0235)	(0.0159)	(0.0359)								
60-64	0.166***	0.245***	0.251***	0.374***								
	(0.0117)	(0.0253)	(0.0204)	(0.0426)								
65-69	0.161***	0.197***	0.262***	0.312***								
	(0.0144)	(0.0351)	(0.0157)	(0.0572)								
70-74	0.132***	0.131***	0.225***	0.215***								
	(0.0123)	(0.0248)	(0.0179)	(0.0411)								
75+	0.0987***	0.158***	0.171***	0.267***								
	(0.00594)	(0.0252)	(0.0118)	(0.0399)								
Gender (ref: female)												
Male	1.003	1.032*	1.003	1.054***	1.024	1.060	0.939***	1.132***	1.039	1.145**	1.059***	1.011
	(0.0105)	(0.0172)	(0.0129)	(0.0172)	(0.0345)	(0.0551)	(0.0213)	(0.0283)	(0.0329)	(0.0699)	(0.0215)	(0.0388)
Birthplace (ref: Native born)												
Other	1.048	0.847**	0.851***	0.719***	0.789***	0.670***	0.836***	0.722***	0.887	0.684***	0.891**	0.784***
	(0.0337)	(0.0703)	(0.0238)	(0.0606)	(0.0536)	(0.0533)	(0.0413)	(0.0693)	(0.0687)	(0.0954)	(0.0488)	(0.0736)
Marital status (ref: Never married/single)												
Married, spouse present	0.924	1.040	1.066	1.191**	1.555***	2.022***	1.090**	1.086	0.958	1.075	0.791***	0.916
	(0.0556)	(0.0838)	(0.0500)	(0.0923)	(0.183)	(0.451)	(0.0413)	(0.0818)	(0.0702)	(0.113)	(0.0470)	(0.114)
Married, spouse absent	1.284***	1.800***	1.257***	1.878***	1.088	2.317***	1.238**	1.584**	1.127	1.541*	1.081	1.856***
	(0.0625)	(0.224)	(0.0616)	(0.252)	(0.247)	(0.505)	(0.115)	(0.308)	(0.102)	(0.400)	(0.118)	(0.425)
Separated	1.805***	1.732***	1.657***	1.644***	1.247	1.230	1.491***	1.468**	1.711***	1.725***	1.443***	1.502***
	(0.0706)	(0.172)	(0.0721)	(0.156)	(0.169)	(0.387)	(0.0959)	(0.249)	(0.157)	(0.311)	(0.132)	(0.222)

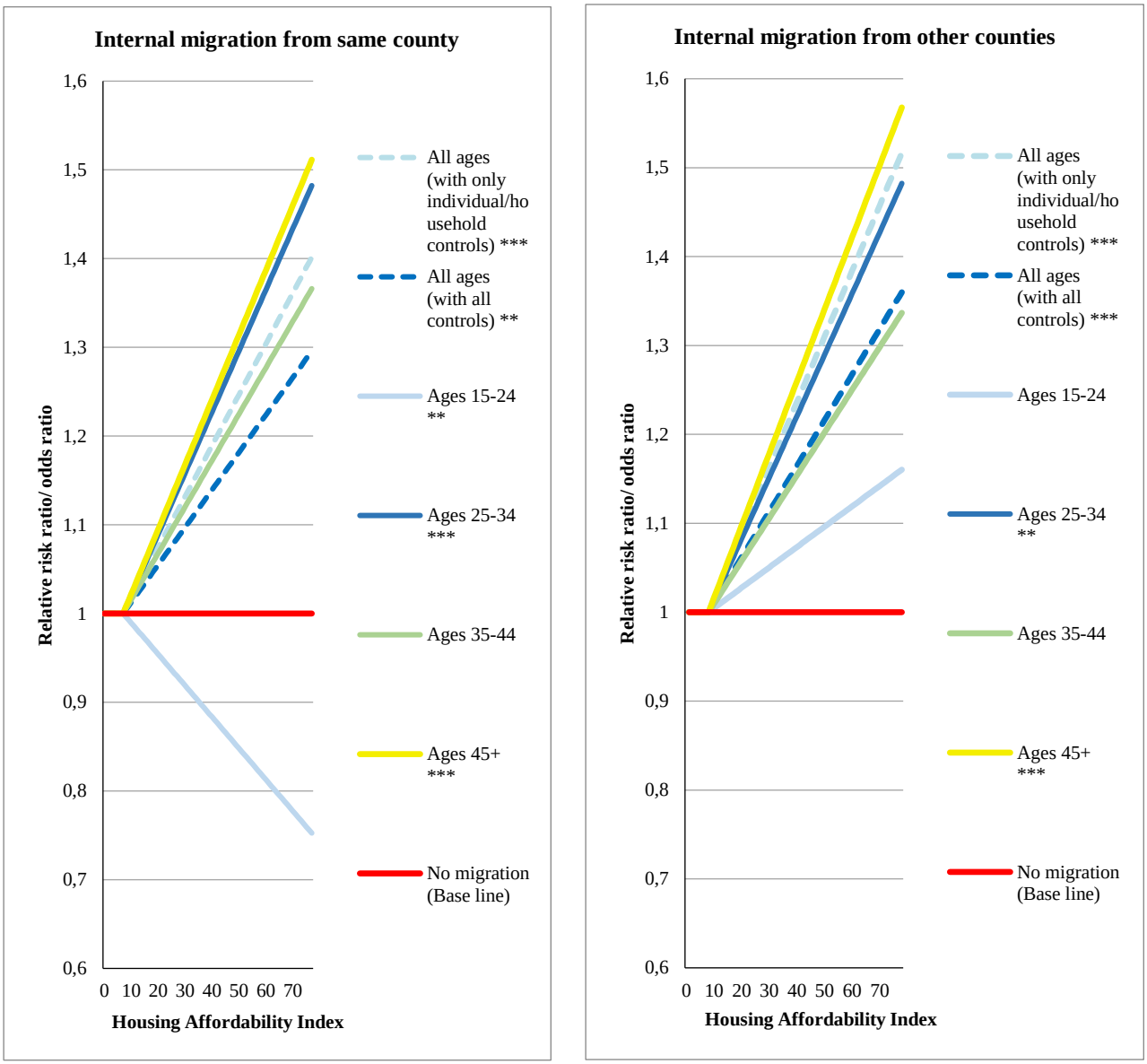
Divorced	1.715***	1.781***	1.526***	1.600***	1.885**	1.638	1.384***	1.660***	1.410***	1.308***	1.274***	1.440***
	(0.0709)	(0.143)	(0.0791)	(0.127)	(0.473)	(0.780)	(0.0756)	(0.236)	(0.0991)	(0.111)	(0.0948)	(0.132)
Widowed	1.414***	1.383***	1.418***	1.428***	0.370	2.562**	1.113	1.105	1.532**	1.052	0.867	1.025
	(0.0812)	(0.166)	(0.104)	(0.181)	(0.404)	(0.951)	(0.256)	(0.503)	(0.267)268)	(0.354)	(0.0895)	(0.154)
Children(ref: None)												
At least one	0.884***	0.637***	0.895***	0.688***	1.296***	0.654**	0.718***	0.668***	0.806***	0.676***	1.143***	0.731***
	(0.0236)	(0.0300)	(0.0173)	(0.0235)	(0.0977)	(0.132)	(0.0319)	(0.0591)	(0.0445)	(0.0496)	(0.0556)	(0.0676)
Race (ref: White)												
Black/negro	1.329***	1.397***	1.064	1.089	1.096	1.197	0.926	0.830	1.167**	1.162**	1.062	1.248**
	(0.0610)	(0.103)	(0.0788)	(0.0829)	(0.153)	(0.132)	(0.0590)	(0.113)	(0.0845)	(0.0850)	(0.0543)	(0.137)
Other	1.006	1.351***	1.102*	1.373***	1.185**	1.304***	0.956	1.161**	1.124	1.525***	1.124	1.616***
	(0.0650)	(0.102)	(0.0550)	(0.0641)	(0.102)	(0.121)	(0.0355)	(0.0798)	(0.0998)	(0.193)	(0.101)	(0.106)
Education (ref: High school or less)												
Middle educated			1.096***	1.349***	1.342***	1.768***	1.081**	1.197***	1.079**	1.373***	1.069	1.292***
			(0.0383)	(0.0998)	(0.126)	(0.191)	(0.0399)	(0.0825)	(0.0389)	(0.153)	(0.0529)	(0.0864)
Higher educated			1.326***	2.040***	2.039***	4.239***	1.303***	2.222***	1.253***	1.834***	1.190***	1.418***
			(0.0681)	(0.173)	(0.151)	(0.865)	(0.0887)	(0.191)	(0.0769)	(0.217)	(0.0752)	(0.107)
Household Income (ref: Central quintile)												
Lowest quintile			1.104***	0.990	1.159***	1.306***	1.166***	0.933	1.087**	0.901	0.976	0.748***
			(0.0291)	(0.0688)	(0.0583)	(0.130)	(0.0555)	(0.0875)	(0.0443)	(0.101)	(0.0566)	(0.0647)
Highest quintile			1.034	1.041	1.015	0.978	0.957	1.021	1.065	1.014	1.163**	1.211**
			(0.0352)	(0.0669)	(0.0641)	(0.0920)	(0.0339)	(0.0850)	(0.0418)	(0.0819)	(0.0846)	(0.111)
Employment status (ref: Employed)												

Unemployed			1.067*	1.592***	0.958	1.168	0.939	1.690***	1.221***	2.027***	1.338***	1.774***
			(0.0389)	(0.113)	(0.0606)	(0.171)	(0.0799)	(0.115)	(0.0860)	(0.383)	(0.0696)	(0.208)
Not in the labor force			0.941*	1.347***	0.810***	1.012	0.949	1.458***	0.904	1.788***	0.908	1.312*
			(0.0317)	(0.109)	(0.0487)	(0.177)	(0.0533)	(0.134)	(0.0603)	(0.233)	(0.0769)	(0.198)
Worked last year (ref: Yes)												
No			0.877***	0.760***	0.785***	0.629***	0.773***	0.742**	1.033	0.902	0.927	0.818
			(0.0191)	(0.0629)	(0.0540)	(0.0941)	(0.0410)	(0.0867)	(0.0955)	(0.127)	(0.0810)	(0.107)
Tenure (ref: Owner)												
Renter			3.555***	3.946***	5.013***	4.462***	2.452***	3.078***	2.469***	3.580***	4.974***	4.452***
			(0.379)	(0.328)	(0.718)	(0.261)	(0.192)	(0.265)	(0.270)	(0.348)	(0.625)	(0.654)
Model summary												
	df	*Chi sq.*	*df*	*Chi sq.*	*df*	*Chi sq.*	*df*	*Chi sq.*	*df*	*Chi sq.*	*df*	*Chi sq.*
Model Wald Chi sq.	29	10,811.90	28	16,173.35	28	4741.32	28	2189.23	29	1786.42	29	3690.47
Pseudo R2		0.0518		0.1009		0.1196		0.0520		0.0575		0.0863

Note: Numbers in parentheses are standard errors. ref.= reference category

*** p < .01, **p < .05, *p < .10 (compared to no migration)

Figure 9 Estimated relative risk ratios/odds ratios for internal migration into California counties by housing affordability for different age groups.



Source: Own elaboration on the data

Note: Figure includes significant and insignificant categories from the models in table 2

*** p < .01, **p < .05, *p < .10 (compared to no migration)

4.3 Estimated internal migration differences by housing affordability from multinomial logistic regression models [Education-specified models]

The third sub-question in this paper asked how the influence between housing affordability and internal migration into California counties was different between various education levels. The other models, therefore, focused on different levels of educational attainment. Three different education groups have been created based on their highest completed level of education. Those who completed a high-school degree or less were referred to as the lesser educated; those who had more education than high-school but less than a four year college degree – that is some years of college or an associate

degree – were referred to as the middle educated; and those with a college degree or higher were seen as the higher educated in this paper. From the results in table 4 it becomes clear that the influence of housing affordability on both categories of internal migration compared to no migration for the middle educated was not significantly different from zero. The lesser educated, on the other hand, showed significant higher relative risks of internal migration into and within California counties when housing affordability increased (consistent with hypothesis 3). When the housing affordability within the destination county increased with one per cent, the relative risk ratio of internal migration from the same county increased with 1.004 (at a 99% confidence interval), while the relative risk ratio of internal migration from another county increased with 1.009 (at a 99.9% confidence interval) for individuals with the lowest completed levels of education. When looking at the higher educated, the influence of housing affordability on internal migration was only significantly different from zero for internal migration from the same county compared to no migration. The relative risks of this internal migration category increased with 1.007 (at a 99% confidence interval) when housing affordability increased with one per cent in their county. When looking at the Wald Chi^2 scores of the specified models for education, again positive scores were noticeable indicating that the predictors have an improvement of fit over a model with no predictors. The Pseudo R^2 of the model for the less educated was 0.0875, while the models for the middle- and higher educated had values just over 0.10. This indicates that housing affordability; in combination with the control variables and the fixed effect explain about 10% of the internal migration variable. This was also the case for the models 1, 2, 3, 4, 5 and 6. The results of the education specified models are shown in the table below. Again, the estimated relative risk ratios/odds ratios of internal migration compared to no migration by housing affordability – but now for the education specified models– were elaborated in a graph, see figure 10. Before explaining the results more extensively, the results of the semi-constructed interviews were first provided in the following subsection. Eventually, the results from both the estimation models and the semi-constructed interviews were explained elaborately in the last part of this section.

Table 3 Estimated Relative Risk Ratios from multinomial logistic regression for internal migration in California counties.

	Model 7 (N=102,847) Lesser educated (High school or less)		Model 8 (N=56,453) Middle educated (More than high school but less than bachelor's degree)		Model 9 (N=54,203) Higher educated (Higher than bachelor's degree)	
Predictor	Internal migration from same county	Internal migration from other county	Internal migration from same county	Internal migration from other county	Internal migration from same county	Internal migration from other county
Housing Affordability Index						
	1.0042**	1.0091***	1.0037	1.0033	1.0071**	1.0049
	(0.00186)	(0.00263)	(0.00328)	(0.00302)	(0.00287)	(0.00395)
Year (ref: 2000-2006)						
2007-2012	0.887*	0.429***	0.809***	0.520***	0.734***	0.602***
	(0.0597)	(0.0497)	(0.0464)	(0.0492)	(0.0583)	(0.0687)
2013-2018	0.624***	0.358***	0.661***	0.453***	0.716***	0.522***
	(0.0373)	(0.0500)	(0.0453)	(0.0309)	(0.0388)	(0.0537)
Age group (ref: 25-29)						
15-19	0.624***	0.536***	0.786**	1.316	0.255**	0.503
	(0.0496)	(0.0623)	(0.0792)	(0.287)	(0.173)	(0.346)
20-24	1.125	1.064	1.101	1.082	1.202***	1.540***

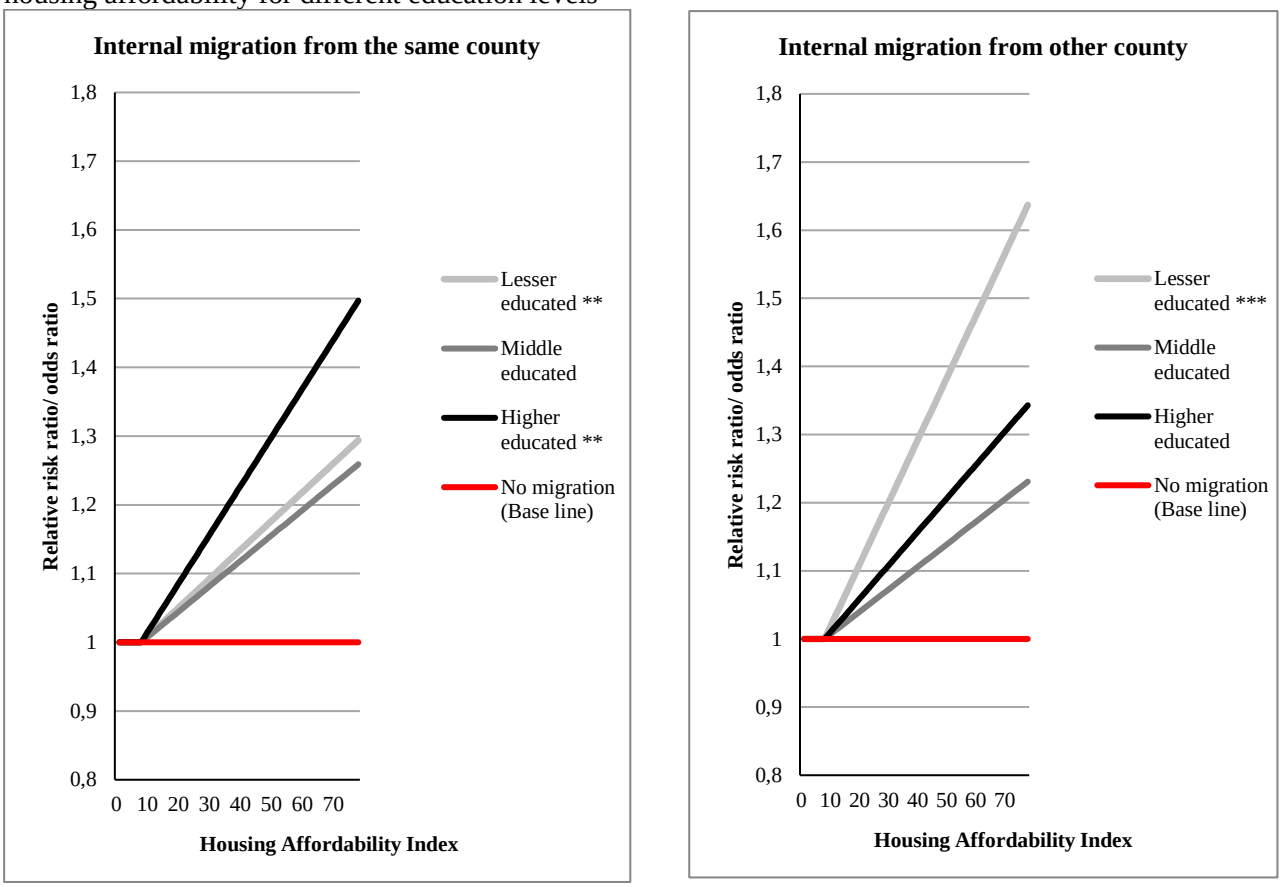
	(0.0948)	(0.0983)	(0.0815)	(0.0911)	(0.0817)	(0.184)
30-34	0.741***	0.795***	0.710***	0.680***	0.692***	0.747***
	(0.0491)	(0.0641)	(0.0302)	(0.0611)	(0.0365)	(0.0633)
35-39	0.589***	0.728***	0.593***	0.615***	0.549***	0.524***
	(0.0559)	(0.0845)	(0.0329)	(0.0556)	(0.0252)	(0.0531)
40-44	0.493***	0.442***	0.495***	0.676***	0.422***	0.402***
	(0.0425)	(0.0593)	(0.0289)	(0.0658)	(0.0239)	(0.0392)
45-49	0.451***	0.436***	0.481***	0.496***	0.356***	0.301***
	(0.0648)	(0.0646)	(0.0332)	(0.0690)	(0.0246)	(0.0427)
50-54	0.419***	0.415***	0.322***	0.465***	0.283***	0.298***
	(0.0514)	(0.0610)	(0.0236)	(0.0677)	(0.0342)	(0.0500)
55-59	0.346***	0.418***	0.316***	0.454***	0.242***	0.304***
	(0.0216)	(0.0717)	(0.0266)	(0.0564)	(0.0294)	(0.0427)
60-64	0.294***	0.353***	0.240***	0.459***	0.199***	0.320***
	(0.0298)	(0.0504)	(0.0414)	(0.0570)	(0.0299)	(0.0458)
65-69	0.302***	0.350***	0.276***	0.330***	0.197***	0.247***
	(0.0189)	(0.0962)	(0.0313)	(0.0886)	(0.0181)	(0.0513)
70-74	0.245***	0.283***	0.279***	0.285***	0.155***	0.106***
	(0.0262)	(0.0618)	(0.0283)	(0.0931)	(0.0301)	(0.0379)
75+	0.188***	0.340***	0.150***	0.292***	0.152***	0.129***
	(0.0186)	(0.0704)	(0.0241)	(0.0607)	(0.0195)	(0.0321)
Gender (ref: female)						
Male	0.984	1.030	1.072	1.048	0.989	1.141***
	(0.0312)	(0.0368)	(0.0473)	(0.0599)	(0.0275)	(0.0378)
Birthplace (ref: Native born)						
Other	0.835***	0.653***	0.832***	0.688***	0.976	0.911
	(0.0347)	(0.0899)	(0.0377)	(0.0675)	(0.0528)	(0.0535)
Marital status (ref: Never married/single)						
Married, spouse present	1.042	1.184	1.136	1.323**	1.081	1.141**
	(0.0432)	(0.132)	(0.110)	(0.159)	(0.0998)	(0.0759)
Married, spouse absent	1.320***	1.726***	1.211*	2.424***	1.180	1.750***
	(0.0875)	(0.332)	(0.121)	(0.412)	(0.125)	(0.281)
Separated	1.470***	1.932***	1.775***	1.560**	2.219***	1.227
	(0.0835)	(0.223)	(0.134)	(0.275)	(0.207)	(0.222)
Divorced	1.555***	1.853***	1.624***	1.511***	1.420***	1.439***
	(0.108)	(0.260)	(0.106)	(0.167)	(0.129)	(0.118)
Widowed	1.365***	1.407**	1.424***	1.415	1.447***	1.280
	(0.126)	(0.218)	(0.160)	(0.336)	(0.165)	(0.284)
Children(ref: None)						
At least one	0.924***	0.680***	0.911**	0.709***	0.867**	0.725***
	(0.0258)	(0.0392)	(0.0387)	(0.0465)	(0.0488)	(0.0668)
Race (ref: White)						

Black/negro	1.217***	1.211	0.938	0.992	0.913	1.016
	(0.0700)	(0.169)	(0.125)	(0.0852)	(0.0525)	(0.105)
Other	1.048	1.601***	1.126**	1.278**	1.010	1.111
	(0.0820)	(0.0928)	(0.0531)	(0.126)	(0.0801)	(0.0797)
Household Income (ref: Central quintile)						
Lowest quintile	1.110***	0.904	1.126***	1.122	1.098***	1.070
	(0.0413)	(0.0825)	(0.0486)	(0.104)	(0.0381)	(0.0603)
Highest quintile	1.045	1.174*	1.052	1.056	0.988	0.971
	(0.0405)	(0.102)	(0.0398)	(0.113)	(0.0584)	(0.0598)
Employment status (ref: Employed)						
Unemployed	1.099	1.349**	1.125**	2.011***	0.871**	1.452***
	(0.0638)	(0.183)	(0.0526)	(0.231)	(0.0611)	(0.173)
Not in the labor force	0.949	1.134	0.930	1.394**	0.919	1.651***
	(0.0546)	(0.188)	(0.0561)	(0.218)	(0.0767)	(0.134)
Worked last year (ref: Yes)						
No	0.888***	0.849	0.871*	0.700**	0.841*	0.755**
	(0.0397)	(0.124)	(0.0617)	(0.100)	(0.0792)	(0.0837)
Tenure (ref: Owner)						
Renter	3.189***	3.416***	4.545***	4.056***	3.067***	4.342***
	(0.370)	(0.399)	(0.607)	(0.354)	(0.280)	(0.364)
Model summary						
	df	*Chi sq.*	*df*	*Chi sq.*	*df*	*Chi sq.*
Model Wald Chi sq.	28	6266.91	28	5477.44	29	4906.96
Pseudo R2		0.0875		0.1203		0.1147

Note: Numbers in parentheses are standard errors. ref.= reference category

*** p < .01, **p < .05, *p < .10 (compared to no migration)

Figure 10 Estimated relative risk ratios/odds ratios for internal migration into California counties by housing affordability for different education levels



Source: Own elaboration on the data

Note: Figure includes significant and insignificant categories from the models in table 3

*** p < .01, **p < .05, *p < .10 (compared to no migration)

4.4 Semi-structured interviews

Even though the emphasis in this paper was placed on providing an estimation of internal migration into California by housing affordability and, thus, relies on quantitative research, the value of qualitative research methods should not be underestimated. Especially for the study in this paper and its related main research question, qualitative instruments could considerably contribute to attain a more comprehensive understanding about housing affordability and how it works on internal migration into California counties. Therefore, one qualitative sub-question was prepared, which was elaborated by conducting two semi-constructed interviews with researches from different research institutions. During the interviews the author of this paper – J.G. Vellinga – was the interviewer. It should be emphasized that the elaborated interviews below are summaries of the original semi-constructed interviews. As explained in section 3.3, semi-constructed interviews can be used as an instrument to explore the data. The aim is not to be representative, but to understand how participants experience issues related to the topic (Valentine, 2005 p.111). Because of that, there was chosen to not include the questions asked in the summaries below but to stay with the essence by only providing the insights from the interviewees. The interview guide including the questions asked can be found in Appendix E. The summaries were structured based upon the topics from the interview guide. In the last part of this subsection, the two interviews were compared and the similarities and differences were described.

Interview with Professor Susan Brown

Results/analysis When the results of the estimation model were submitted to Professor Susan Brown, she emphasized that the results made a lot of sense. Professor Brown: "Lesser skilled and younger people should be leaving the poorest areas because there is nothing there for them, but they should also not be going to counties like Santa Barbara because there it is really expensive for them. Instead they should move to affordable counties with jobs and amenities. Besides, commuting in California is on the uptick, but there is a limit to commuting. And so the counties that are nearer to the large metropolitan areas are growing the fastest. Actually, there is a huge pull towards these more exurban or suburban counties because they are cheaper. For example areas near Greater Sacramento and The Inland Empire; all of these areas are growing. However, the counties that are really cheap up in Northern California – the rural ones – are not that popular because of weaker labor markets and people can't commute from there. Therefore, the results would make a lot of sense, since they are limited to counties with a certain proximity to a metropolitan area."

Housing affordability in California Professor Brown: "At the moment, housing affordability is really bad in some California counties. Prices have gone up everywhere but there is a huge variation in the size of the increase, and California is particularly high right now. Orange County, Los Angeles County and San Francisco County are examples of counties where housing is rather unaffordable at this moment. It also means that if lower income people want to stay in these areas they will have to be much more flexible and creative on living arrangements than they probably ever thought needing to be. More people are living together in in-law apartments and with flexible arrangements that were more common 100 year ago than for recent generations. People just have to do that, because right now as of the affordability crisis – which could change, but is not going to change anytime soon – building materials and land cost aren't going to go down that much. There is no reason to think they would. The attractive and expensive counties might get more housing on the market and that would ease up the demand a little bit, but it is going to be reasonably expensive. So yes, there is a really big problem. Many people have moved away from the expensive counties, not only because it is cheaper to live in another county, but also because they could then have more space. For example, for families it would make a lot more sense to purchase a house with a yard than just living in a two-bedroom apartment. This is much more conducive for young families. So, regarding the willingness to pay: Apartments are going to seem very small if you get older. But there are trade-offs. It is all a series of trade-offs and although housing affordability is a key one, many people are willing to pay for and live in some smaller house or apartment with closer proximity to the urban area. Especially when they are convinced that living in these places is a good investment. There are some that fear the housing market due to the recent financial crisis. Especially, in the cheaper areas in California – that were hit the hardest."

The cause Professor Brown: "The main cause of the unaffordability crisis in some counties is limited amounts of developable land. There just isn't room to develop without becoming much denser. And when it becomes much denser there are many people objecting to that because of traffic increases. For example, Los Angeles County – which is surrounded by mountains – just approved a new 19.000 house development project in the northern part of the county, somewhat closer to Bakersfield. It is high, flat, and grazing ground, but it is a high fire risk area. Developers are saying: 'Well we are going to do our best to put in housing and landscape that is fire resistant.' But there are all kinds of people saying: 'Are you nuts?!' And yet that is the last really big development ground in the county. These are the kinds of issues that people are grappling with. 'Do you create desperately needed new housing and try to make it as fire proof as is possible?' 'Do you put in more high rises in areas that do not really have the infrastructure for traffic and do not have the money for public transportation?' Besides it is hard to create a whole new lifestyle trying to Manhattanize Los Angeles. So it is a real problem and it all comes down to a lack of housing. The cost of land therefore goes way up. It is a little bit more expensive to build in in California than in other states, but the building costs are not really the issue, it is rather the land cost. Besides, homeowners also have few incentives to sell and leave their homes. For example, the proposition 13 pass back in 1977, ensured that homeowners could keep their property taxes at the same level for all future years at the moment when they bought the house. So, if

they would move and purchase a new house, their property taxes would go up. Because of that homeowners are less likely to move in California once they bought."

The affected ones Professor Brown continued: "Yet, housing affordability is absolutely important for the counties of California. Because if people cannot afford to live here, they won't come. It is clearly the less skilled – that is the lesser educated – that are affected and it is starting to affect the more skilled – those with higher educations. The higher skilled have more money and they can handle this. But there comes a point when they say: "I can take my skills, move them elsewhere and, instead, get a lot more house". So it is starting to affect them as well. Besides, it does make sense that young adults would migrate into the more suburban or exurban counties relatively close to urban areas. However, even though telecommuting is on the uprise, some of these migrants would have to commute in order to access their jobs. Yet, the young adults are extremely important for California counties. They are the ones who get to the age of having children; the ones whose children go to schools; and they are the consumers who have to buy furniture, baby stuff and they are the ones who can fill the demand on the labor market. All in all, they are the ones who are starting out. The housing affordability crisis can have different consequences for the young adults and the lesser skilled: 1. They might move out from California to other states where housing affordability is not that bad; 2. They might try to commute and live somewhere further from the urban and expensive counties; 3. They can live much more densely within the urban counties. With regard to the last alternative, people have to be willing to trade in space for easy access to jobs and amenities. Theoretically, people don't need a single person bathroom. They can get communal housing with shared kitchens which are less expensive. The result of more densely living is that young people postpone getting their own apartment and are doubling up, but there is only a fine amount of time where they can do that because at some point they want to get married, have a family or are tired of living like this. This is a very severe problem and there is some evidence when looking at living arrangements in California which indicates that the proportion of people who live alone and are young in California is way lower than it is in other states."

The balance Professor Brown: "Because housing affordability is so important and because it can be solved to some extent by generating more housing supply, local policymakers would have to build denser apartments, and in order to get away with denser apartments, they will have to build wider streets. Irvine is a good example that is trying, by having a lot of schools and shopping centers close to housing areas. So people don't have to go far in order to buy something. They can shop and live in relatively contained areas. So it doesn't require a lot of big trips. Even with that they have hugely growing traffic. An example comes from Newport Beach, Orange County who have tried to build high skyrise buildings. This has cost a huge backlash as Newport Beach proposed building enormous apartments of 25 stories high, but the city council that proposed it got voted out, because people didn't want that kind of density coming in without any improvements of roads. So on the one hand affordable housing is necessary to keep and attract people, but on the other hand building affordable housing comes with many other problems such as traffic congestion. Policymakers should try to find a proper balance between building affordable housing and controlling the problems resulting from that. But the perfect balance? I wish I had the answer for you."

Interview with Doctor Wallace Walrod

Results/analysis When sharing the results with Doctor Wallace Walrod he thought the analysis was really good and interesting. However, he also recommended looking separately at the time frames within the 2000-2018 year period, because he expected that the migration of young adults to cheaper counties mainly used to happen in the beginning of this period – before the financial crisis. Doctor Walrod: "Although housing affordability might still impact internal migration into California to some extent, the labor markets – especially the tech based labor market – work as huge pull-factors for internal migration into California counties. Evidence comes from Orange County that even though it got more expensive in 2017, outmigration and millennial migration actually started to fall a bit. Most likely this results from the robust job markets in Orange County, but also in – for example – the Bay Area and Silicon Valley. At some point there will indeed be a tipping point regarding housing affordability and the choice where to migrate but people sometimes are just so attracted to being in

particular places and industries that they won't stop moving there. Another interesting way of looking at the analysis is by pulling apart the counties and looking at the influence of housing affordability on the migration into single counties over time. For example, looking at the influence in Silicon Valley by pulling out Santa Clara County. Even though Silicon Valley is spread over several counties, Santa Clara is the hearth of it. That would be something for future research."

Housing affordability in California Doctor Walrod: "So, actually houses are not quite as unaffordable as they were during the bubble of 2005/2006. It is getting close but there are some signs that housing price appreciation is starting to slow a little bit. Some econometric models even show that housing price appreciation is continuing to slow for the next years. Besides – when coming out of the recession – one of the biggest things that puzzled people was that even when unemployment rates were going down, wage levels did not really rise correspondingly. Just in the last year - both because of the market and also because of laws related to minimum wage increases – people started to see wage increases. Therefore, it is most likely that the expensive counties will probably hit the low in the affordability cycle this year or even last year. Because of housing affordability is also a measure of income and house price appreciation starts to slow, it is expected that houses eventually become more affordable again. The decision where to live is almost like a whole sociology. The housing affordability mechanism and the related decision where to migrate is primarily based upon the price of a house and the type of unit being built. As mentioned before, the price of a house or apartment really depends on the location and on the available supply within the county. The closer to urbanization, the higher the price. Besides, environmental regulation in California does add significant cost to a house, which can add up to a couple of hundred thousand dollars per housing unit. Even though this would not be geographically specific to a particular county, because it is primarily a state policy, it is something to keep in mind. With regard to the type of unit being build there is some geographic variation. Compared to the mid-nineteenth century, much larger housing units are being built nowadays. That is because for the most part California has become wealthier as a state and the people who can afford it want more space. So developers respond to this market demands. However, there is emerging a market for smaller housing units, especially in areas that are highly dense and offer access to other kinds of amenities. The young adults and lesser skilled workers in specific are interested in these kinds of housing units. However, it is not as easy to propose that and get it through the process, because people are not familiar with that type of product, so it is often voted down. Nevertheless, some developers are starting to move into that type of product. This could contribute to solve many problems, including the housing affordability issue."

The Cause Doctor Walrod: "Housing affordability rates in California are primarily driven up because of the lack of supply in the coastal counties. Even though these are the regions with the most robust job markets, they are not the ones who are building the houses that are needed because of a) there being not enough land and b) community opposition. Although the market is starting to accept higher density building, it is the city councils, planning commissions and residents who oppose it. This is one of the big disadvantages of land use being regulated locally in California. However, these issues are usually not the case for the more suburban or exurban counties such as Riverside and San Bernardino, who have a lot more land and are more willing to build houses. For these counties it is not (yet) necessary to build high-density. Housing markets are pretty direct demand-supply markets; almost a pure supply-demand issue. So, obviously price appreciation occurs when there is more demand than supply. There have been a couple of factors that have made housing even more expensive in the last few years. For Orange County probably the biggest one is international demand. There are a lot of foreign buyers that would come in here who sometimes do not even live in the house. The other thing is – the other side of the equation – that homeownership levels are much lower than they have been for a long time. So, what happened is that coming out of the recession investment groups bought op foreclosed properties, and they are renting them out now. At some point they will probably sell them, but there are a lot of single family houses now that are being rented. At the same time there is also a shortage of apartment units, and those prices have gone up tremendously as well. Probably those got even less affordable than normal houses. To contextualize all of this, the biggest issue is just the lack of supply. Yet, when looking at it from a labor market perspective, housing is the infrastructure counties need in order to house workers. Besides, expensive counties such Orange County are rapidly aging, which reduces the number of workers per house, and even more reducing supply on the housing

market. Housing unaffordability is probably the number one issue that can hold California back, because California has always been a place that attracted people from other places. When looking at the labor market supply in the entertainment and high tech industries in California counties, it was almost always filled with people from other places who migrated into the state. Assuming the local economies within California continue to grow, there will be more jobs and thus higher labor market demand in the upcoming years. Even though there might be recessions here and there over time, Orange County – for example – is expected to grow between 2 and 2.5% in terms of jobs. However, when younger people decide to move into more affordable counties, the urban counties could have not enough workers to fulfill the demand on the labor market, and so businesses could consider not expanding here but rather at those more affordable counties. There is some evidence that this already started to happen in some cases. For example, instead of expanding in Orange County, there are some businesses that started to invest in Riverside and San Bernardino County. So in the end, jobs follow the people, and it already started to happen."

The affected ones Doctor Walrod: "Those who have less income to spend are those mostly affected by housing affordability issues. However, that is a problem since those people are still important for counties. When looking at age, it is definitely recent graduates and younger workers who are really important for future economic prosperity in California counties. Losing this age group also implies that counties lose their children. A surprising statistic is that a place like Orange County is going to lose 10% of the population of aged 5 to 17 over the next ten years, because the county has lost their fathers and mothers. That makes some hard choices of having to close schools but also on the workforce perspective. Labor demand is expected to keep growing, and when there is not enough supply younger people have to be attracted somehow to the counties. Concerning education, it is not only the higher educated who are important for California counties. It is expected that the demand for middle-skilled workers will grow significantly in the next years. The labor market of many California counties really moved to a skill-based one which does not necessarily require a four year college degree. However, most jobs do require an above high school degree. The people that really get left behind here are those who did not complete high school or completed high school and stopped there. Even though service workers are needed to facilitate the needs of the higher skilled workers, their incomes are generally not likely to cover up living expenses in the expensive areas."

The balance Doctor Walrod: "Increasing housing supply is the easiest way to make housing more affordable again. Everything else gets more complicated and difficult when not increasing supply in the housing market. Besides, flexibility with the type of product and being open to that is important, as mentioned before. Generating more housing supply indeed goes with infrastructural issues, but counties can take care of that as long as there is good planning. A real opportunity to find a balance between affordable housing and the number of people living in a county and the related congestion problems to that is the mixed use of land."

Similarities and differences

Even though Prof. Brown and Dr. Walrod have different research interests – academic versus business – both researchers agree that housing unaffordability is one of the biggest issues that can California hold back in the next years. It is primarily the coastal counties with the strongest labor markets that are becoming less and less affordable. There was also agreement when they both talked about the main cause of the low housing affordability rates in some of the state's counties – especially the expensive coastal counties: The lack of supply. NIMBY-behavior of citizens and local representatives is an important factor causing restrictions related to generating housing supply by hindering the construction of higher density housing units that could ease up the demand. Besides, homeowners have many incentives to stay in their houses – also known as the homeowner lock – which also limits the housing supply in some counties. However, the two researchers also have different perspectives of looking at this problem. Prof. Brown thinks that the housing affordability crisis is not going to change anytime soon. Additional supply in the expensive counties is going to be reasonable expensive, and therefore less affluent residents are likely having to change their living arrangements by either moving out of – or not moving into these counties, or living much more densely, flexible and creative in the

expensive counties. The affected ones are probably those with the least spendable income on housing; so it would make sense that the young adults and the lesser skilled are those who are mostly affected by housing affordability in California, and are those who most likely will migrate into more affordable counties. However, the professor also emphasizes that it is always a trade-off where individuals have to choose between space and easy access to jobs and amenities. Dr. Walrod – on the contrary – expects that housing affordability is at an all-time-low at this moment. Evidence shows that housing price appreciation will start to slow, while wages started to increase. Therefore, he expects that houses eventually will become more affordable again, also in the expensive counties. However, even though the coastal counties are expensive right now, these are also the places where people want to be. Although housing affordability might still impact internal migration into California to some extent, the labor markets – especially the tech based labor market – work as huge pull-factors for internal migration into California counties. Since people are so attracted to being in particular places and industries, it is unlikely that they will stop moving there. In line with prof. Brown, Dr. Walrod expects that the lower income people are the ones who are mostly affected by housing unaffordability in the expensive California counties. Those who really get left behind are the ones who did not complete high school or complete high school and stop there. Both researchers agree that the solution of the housing affordability crisis lies in increasing housing supply, but also implies that a proper balance should be found between increasing the supply and planning ahead on the problems resulting from that – e.g. congestion and nuisance.

4.5 Discussion of the results

Even though the wage levels did not rise correspondently with declining unemployment rates, it were not the wage levels that have led to low housing affordability rates in some of California's counties. From the semi-constructed interviews it became clear that even though it is a little bit more expensive to build in California, it were not the building costs that drove up the expenses of houses but rather the expensive prices for land made some counties really unaffordable. According to Alonso's Theory of Urban Land Market (1964) land prices are not just the result of market mechanism – that is, supply and demand for housing – but more important, they are the result of competition for space, where the price of land follows a series of bid rent curves in which land prices decrease when it is located further away from the urban center. Individuals choose their location by finding the place that satisfies their space needs and location preferences with their budget constraint (Alonso, 1964; Dieleman & Wegener, 2004; Von Thünen, 1826). Historically, this implied that the more affluent folks desired to live further away from the 'noisy and dirty' city centers since they were able to pay for the transportation costs of getting there, whereas the lower income chose their location in high-density housing areas near the urban center because they could not afford to exchange land for other factors and were not able to pay for the transportation costs to get to the city. However, over the past few decades also the moneyed ones started to move back to the urban centers. According to Florida (2017, p.21) urban places have the highest variety of amenities nowadays, and that's why these places have a huge influence on the individual's decision *where to migrate.* Besides, most knowledge-based, high-tech industries – which shape California's strongest economic asset right now and attract many internal migrants – are primarily located in the coastal urban counties of California. Historically, the affluent people could easily access the urban areas from further away. Nowadays, that has become way more difficult, partly because of factors such as urban sprawl and congestion. Figure 2a and figure 2b illustrated how residential land allocation was divided among three different income groups with preferences respectively for land consumption (historically) and accessibility (currently). However, the "back-to-the-city movement" – as cited by Florida (2017, p.58) – also implied a fierce competition for urban land in some of California's counties. As described in the second section of this paper: If more people want to cluster in (urban) space; the more competition there will be; the more expensive land will get; and the higher house prices will become. As a result, housing affordability rates have mostly dropped in the attractive urban – primarily coastal – counties of California, and it were those counties where houses were least affordable. However, the study in this paper showed that housing affordability had significant influences on the individual's decision to either migrate within

the boundaries of their county or to migrate by even crossing their county borders, which may have major implications related to the state's economy.

The first two multinomial logistic regression models in section 4.2 – prepared for all societal groups in California – showed significant increased risks of internal migration into a California county from the same county and from other counties compared to no migration by an increase in the county's housing affordability (consistent with hypothesis 1). This implies that the individual's decision to migrate into California counties is significantly influenced by housing affordability. The findings of the interviews indicated that many people have moved away from the expensive counties, not only because it is cheaper to live in other counties, but also because they can offer more space. The locational trade-off is between land consumption and the distance to the urban center. The less urban counties are growing because they are cheaper, and also increasingly start to offer more and better amenities. Yet, it is the coastal urban counties which create the best paid, high tech jobs which are also expected to work as a major pull factor of internal migration into these counties. The economies in these counties shape the future for the state of California, and since their natural population is not likely to meet the increasing demand on the labor markets, these counties will mostly rely on internal migration into these counties from younger people and people with different educational backgrounds. However, the results of the age- and education specified models indicated that the young adults (aged twenty-five to thirty-four) and least educated were the only ones whose relative risks of internal migration from other counties into a California county were significantly influenced by housing affordability. In both cases these relative risks increased when housing affordability in the destination county increased, as table 3 and 4, and figure 9 and 10 indicate. These results support the second and third hypothesis arguing that the relationship between housing affordability and internal migration into California was strongest for respectively the young adults (aged twenty-five to thirty-four) and the lesser educated. People aged between twenty and thirty-four – as illustrated in figure 8 – generally have the highest internal migration propensities. It is the young adults who experience the most life- and career-cycle events which influence their decision *to migrate*. Family expansion, marriage, buying a first house, working on a first job or job promotion are all examples of changes in the life- and career cycle that frequently apply to the 'young-adult' age-group (Franklin, 2003; Roseman, 1977). Now it also becomes clear that the young adults are significantly influenced by regional housing affordability, which influences their decision *where to migrate.* The same goes for the least educated individuals in the analysis. They are generally the ones with the least spendable income on housing. From economic location theory this implies that they would be the ones driven out of the urban centers by higher income people or not able to win the competition for urban space. As a result the study in this research proves that their decision *where to migrate* in California was significantly influenced by housing affordability. As Professor Brown described in the first semi-constructed interview the housing affordability crisis can have different consequences for the young adults and the lesser skilled: They might move out from California to other states where housing affordability is not that bad; they might try to commute and live somewhere further from the urban and expensive counties; or they can live much more densely within the urban counties. The second consequence seems to be true in the quantitative study of this research.

Professor Brown also expected that housing affordability is starting to affect the higher educated. Only the results from the highest educated who migrated within the same county support this claim. Figure 10 shows that they even seemed to have the highest significant increased relative risks of internal migration into the same county compared to no migration when housing affordability increased. Remarkable was that the oldest age group (aged forty-five and over) showed similar results with the highest educated group. Compared with all other age groups, the oldest age group had the highest significant increased relative risks of internal migration within the same county compared to no migration when housing affordability increased in the analysis, as shown in figure 9. From Alonso's Theory of Urban Land Market this would make sense, since the higher educated and the older adults are generally the ones who can generally bid the highest price for land. So, they will move and bid for places within their county when these get cheaper, but won't move to other counties because of weaker labor markets and lower quality amenities. Besides, life- and career cycle events explain why people with older ages are generally the ones who are looking for more affordable houses because – for example – their children leave the household, they retire or they lose a spouse (Roseman, 1977;

Zaidi, 2014). Life and career events related to older ages can trigger them to find a cheaper – often also a smaller – house somewhere in the same county. When houses get more affordable that would be the perfect time to move. It should also be emphasized that recent competition for urban land also brings something else about, namely the competition from the super-rich and investors (Freeman, 2012 p.188-229). Because of that, it would make sense that the higher educated and the older age groups are willing to move within the county when houses get more affordable. They will just buy a more affordable house and sell their own house to those that are willing to pay the price. Another remarkable result was that the relative risks of migration within the county decreased for the fifteen to twenty-four years old age group when housing affordability increased. This would imply that they were less likely to migrate when housing affordability increased in the county. For this age group other mechanisms – such as still going to school or being part of their parental household – may cause deflections in their internal migration decision by housing affordability. However, previous research does not provide any evidence supporting this finding. For the ages thirty-five to forty-four no significant results were found regarding their migration behavior by housing affordability. Housing affordability also did not significantly influence migration behavior of the middle-skilled people – those who completed a higher degree than high-school, but less than a four year college degree. From a labour market perspective it can be argued that the demand for these societal groups is high and is even expected to rise in the next years. Besides middle aged and -educated people usually are able to bid higher prices for land than younger and less educated ones. Therefore, they are less likely to be priced out from urban areas. The results from the statistical analysis in this research indeed show that more affordable housing does not significantly gives them incentives to either move to other counties or move within the county.

Four sub-questions were prepared in order to answer the main question in this research: What is the influence of housing affordability on internal migration into California counties? The first sub-question asked to what extent housing affordability influences internal migration into California counties. It became clear that the first two models showed that the relative risk ratios of internal migration into a California county from the same county and from other counties compared to no migration significantly increased when the housing affordability index rate in the destination county increased. This supports the first hypothesis mentioned in section 2.5. The second and third sub-questions asked how the relationship between housing affordability and internal migration into California counties was different between age- and education groups. The age- and education specified models indicated that only groups in which individual internal migration behavior to other counties was significantly influenced by higher housing affordability in those counties were the young adults (aged twenty five to thirty-four) and the least (or lesser) educated. When housing affordability increased in the destination county, the relative risks of internal migration versus no migration increased for individuals in these groups. However, when looking at internal migration from the same county compared to no migration, it became clear that the oldest age group (aged forty-five and over) had significantly higher increased relative risks than the young adults and the age group fifteen to twenty-four when housing affordability within the county increased. The results applied to the higher skilled compared to the lesser skilled. So, the second and third hypotheses – prepared in section 2.5 – only hold when looking at internal migration to other counties. The fourth and last sub-question asked how housing affordability influences internal migration into California counties by using semi-constructed interviews. The results from these interviews were used to explain the results from the quantitative analysis in this section and to understand the wider context in which housing affordability influences internal migration in the state. Housing affordability clearly influences internal migration into the state, and its influence impacts the individual's decision *where to migrate* – or in which county to migrate – for different age- and education groups.

Section 5 – Conclusion

5.1 Conclusions

This paper is one of the several recent contributions to literature about the housing affordability and internal migration in California. Previous papers have focused on internal migrants moving out from the state, who move in patterns which imply that housing affordability is a driving force behind these moves. These concerns are also part of the present paper. Additionally, this research builds on those concerns and expected that when there are migration patterns *from* the state by housing affordability, there would also be a migration pattern *into* the state. Another new contribution was that the study in this paper examined individual migration choices by housing affordability, and can therefore be seen as an elaboration on the more recent internal migration literature that tries to place more emphasis on obtaining a more sophisticated individual background on why people migrate and the underlying mechanisms – in specific this paper elaborates on the individual's decision *where to migrate*. Understanding the role that housing affordability has on internal migration into California could be useful for either facilitating or counteracting on residential choice. Undesirable mobility can cause or worsen regional problems, and can obstruct future economic development. Both quantitative and qualitative research methods were used to determine the influence of internal migration by housing affordability in California counties. While semi-constructed interviews where used to obtain a more comprehensive understanding of housing affordability and internal migration in California, multinomial logit frameworks were used to estimate the individual's relative risk ratios of internal migration – either from the same county or from other counties – compared to no migration by housing affordability. The results from the first two models indicated that increases of housing affordability rates in the destination county significantly increased the relative risks of both kinds of internal migration compared to no migration into that county. The multinomial logistic models were also specified for different age- and education groups. The results from these models clearly indicate that the least advantaged and most economically vulnerable residents were the ones whose internal migration behavior was (significantly) positively influenced by housing affordability into new or other California counties, while individuals from the more advantaged groups only changed their location within the same county or did not change their location at all when housing affordability increased. Yet, the most economically vulnerable will be of great importance for future economic prosperity in California, which is likely to be shaped by the strong labour markets in the low affordable coastal urban counties. However, the solution is not to shut down the process of wealth creation in these counties, but to make their flourishing economies more encompassing and inclusive. It would not make sense to discourage investments in cities and urban neighborhoods in which houses have become highly unaffordable. Rather than stopping the market forces that have led to the economic revitalization of multiple urban counties but also drove up the housing prices, urban policy should improve housing options for those who tend to migrate to more affordable counties, not just by increasing supply but by increasing affordable supply in the counties where the strong labour markets which will shape California's future are easy to access. However, an important contribution of the paper is to show that the results do not simply replicate previous work, rather it provides a new understanding of how the individual's decision *where to migrate* can be influenced by regional housing affordability.

5.2 Recommendations for future research

Even though many other predictor variables were used to control for confounding variables, the results from the study in this paper did not elaborate extensively on these other predictors. Yet, these other variables could also offer important insights into the internal migration preferences of individuals. As Dr. Walrod recommended in the second semi-constructed interview, looking separately at 2000-2018 time frames could provide valuable insights related to the influence of housing affordability on internal migration into California. Even though housing affordability might still influence internal migration into California to some extent, he expected that changing migration

preferences by housing affordability was something that used to be, as the strong high-tech labor markets work as huge pull factors for internal migration into California counties. The quantitative study in this paper provided an estimation of the influence of housing affordability on internal migration into the counties that were studied. Another interesting way of looking at the analysis would be by pulling apart the counties and looking at the influence of housing affordability on the migration into single counties over time. As the study in this paper focused on internal migration patterns into counties, it should also be noted that internal migration can also happen at smaller geographical scale levels. In this research it became clear that housing is more affordable in the less urban counties of California and those are the places where the lesser educated and the young adults tend to move to. Yet, at smaller geographical scale levels there are also enormous differences between housing affordability rates. When looking at the city of Los Angeles, for example, the more affluent people have occupied houses and land along the coastline, around the urban core, universities and knowledge institutions, while being surrounded by the less advantaged ones who live in cheaper zones. Interesting would be to understand how housing affordability works on internal migration within these smaller geographical units. Despite the fact that the study in this paper offers opportunities for further research, the results from the multinomial logit frameworks and semi-constructed interviews provide valuable information about housing affordability in California counties and its influence on internal migration patterns in(to) the state. It is the less advantaged groups who were hit the hardest by low affordability of houses and who tend to move to more affordable counties in California. Increasing the supply of affordable housing seems to be the only solution to solve the housing affordability crisis in the expensive coastal urban counties. Yet, the question remains: How to increase the supply of affordable housing in these counties without attaining the problems that emerge as a consequence of that?

Reference list

- Allison, P. D. (2009). *Fixed effects regression models* (Vol. 160). Thousand Oaks: SAGE publications.
- Alonso, W. (1964). *Location and Land Use. Cambridge*, MA: Harvard University Press.
- Barro, R. J. (1991). Economic growth in a cross section of countries. *The quarterly journal of economics*, 106(2), 407-443.
- Baum-Snow, N., & Hartley, D. A. (2016). *Accounting for central neighborhood change, 1980-2010.* FRB of Chicago Working Paper No. WP-2016-9. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2839686## [Accessed 15 January 2019}
- Bean, F., Brown, S., & Pullés, S. (2018). Migration and the California Dream: Past, present and future. Irvine: University of California, Irvine[5]
- Benetsky, M. J., & Fields, A. (2015). *Millennial migration: How has the Great Recession affected the migration of a cohort as it came of age?* US Census Bureau, SEHSD Working Paper, 1. Available at: https://census.gov/content/dam/Census/library/working-papers/2014/demo/MillennialMigration_Benetsky-Fields.pdf [Accessed 20 November 2018]
- Blomquist, G. C., Berger, M. C., & Hoehn, J. P. (1988). New estimates of quality of life in urban areas. *The American Economic Review*, 89-107.

[5] This is a draft of a paper presented at the Ninth Annual University of California International Migration Conference on ―Immigrant California: Policies and Politics,‖ sponsored by the California Immigration Research Initiative (CIRI), University of California, San Diego, March 2, 2018. Available on request.

- Bricker, J., & Bucks, B. (2016). Negative home equity, economic insecurity, and household mobility over the Great Recession. *Journal of Urban Economics*, *91*, 1-12.
- California Association of Realtors (2018a). Housing Affordability Index – Traditional. *California Association of Realtors*. Available at: https://www.car.org/marketdata/data/haitraditional/ [Accessed 9 July 2018]
- California Association of Realtors (2018b). Housing Affordability Index – Traditional Methodology. *California Association of Realtors*. Available at: https://www.car.org/en/marketdata/data/haimethodology [Accessed 10 September 2018]
- Gallagher, L. (2014). *The end of the suburbs: Where the American dream is moving.* New York: Portfolio Penguin.
- Cameron, J. (2005). Focusing on the focus group. *Qualitative research methods in human geography*, 2, 116-132.
- Carr, L. T. (1994). The strengths and weaknesses of quantitative and qualitative research: what method for nursing? *Journal of advanced nursing*, 20(4), 716-721.
- Clark, W. A., & Lisowski, W. (2017). Decisions to move and decisions to stay: Life course events and mobility outcomes. *Housing Studies*, 32(5), 547-565.
- Cohen, D. (2011). *Braceros: Migrant citizens and transnational subjects in the postwar United States and Mexico.* Chapel Hill: University of North Carolina Press.
- Cooke, T. J. (2013). Internal migration in decline. *The Professional Geographer*, 65(4), 664-675.
- Davies, H. T. O., Crombie, I. K., & Tavakoli, M. (1998). When can odds ratios mislead? *Bmj*, 316(7136), 989-991.
- Dieleman, F., & Wegener, M. (2004). Compact city and urban sprawl. *Built Environment*, 30(4), 308-323.
- Ewing, R., Hamidi, S., Grace, J. B., & Wei, Y. D. (2016). Does urban sprawl hold down upward mobility? *Landscape and Urban Planning*, 148, 80-88.
- Flood, S., King, M., Ruggles, S., and Warren, J.R. Integrated Public Use Microdata Series, Current Population Survey: Version 6.0 [dataset]. Minneapolis, MN: *IPUMS, 2018*. Available at: https://doi.org/10.18128/D030.V6.0 [Accessed 17 October 2018]
- Florida, R. (2002). Bohemia and economic geography. *Journal of economic geography*, 2(1), 55-71.
- Florida, R. (2017). *The New Urban Crisis: Gentrification, housing bubbles, growing inequality, and what we can do about it*. New York: Basic Books.
- Florida, R., & King, K. (2016). Rise of the global startup city: the geography of venture capital investment in cities and metros across the globe. *Martin Prosperity Institute.* Available at: http://martinprosperity.org/media/Rise-of-the-Global-Startup-City.pdf [Accessed on 07 August 2018]
- Florida, R., & Mellander, C. (2015). Segregated city: The geography of economic segregation in America's metros. *Martin Prosperity Institute*. Available at: http://martinprosperity.org/media/Segregated%20City.pdf [Accessed 15 January 2019]
- Florida, R., & Mellander, C. (2016). The Geography of the Global Super Rich. *Martin Prosperity Institute*. Available at: http://martinprosperity.org/media/2016-MPIWP-004_The-Geography-of-the-Global-Super-Rich_Academic_Florida-Mellander.pdf [Accessed 15 January 2019]
- Franklin, R. S. (2003). Migration of the young, single, and college educated: 1995 to 2000. *US Department of Commerce, Economics and Statistics Administration, US Census*

Bureau. Available at: https://usa.ipums.org/usa/resources/voliii/pubdocs/2000/censr-12.pdf [Accessed 20 November 2018]

- Freeland, C. (2012). *Plutocrats: The rise of the new global super-rich and the fall of everyone else*. New York: Penguin.
- Frey, W. H. (1995). Immigration and internal migration "flight": A California case study. *Population and Environment*, 16(4), 353-375.
- Frey, W. H. (2014). Will this be the decade of big city growth*? The Brookings Institution*. Available at: https://www.brookings.edu/opinions/will-this-be-the-decade-of-big-city-growth/ [Accessed 16 January 2019]
- Frey, W. H. (2016). Demographic Reversal: Cities Thrive, Suburbs Sputter. *The Brookings Institution*. Available at: https://www.brookings.edu/opinions/demographic-reversal-cities-thrive-suburbs-sputter/ [Accessed 16 January 2019]
- Fuller, C. (2016). Rockin' the Suburbs: Home Values and Rents in Urban, Suburban and Rural Areas. *Zillow Research.* Available at: https://www.zillow.com/research/urban-suburban-rural-values-rents-11714/ [Accessed 16 January 2019]
- Fuller, J., & Raman, M. (2016). Bridge the gap: Rebuiliding America's Middle Skills. *Harvard Business School*. Available at: https://www.hbs.edu/competitiveness/Documents/bridge-the-gap.pdf [Accessed 22 December 2018]
- Glaeser, E. L. (1998). Are cities dying? *Journal of economic perspectives*, 12(2), 139-160.
- Glaeser, E. L., & Mare, D. C. (2001). Cities and skills. *Journal of labor economics*, 19(2), 316-342.
- Glasmeier, A. K. (2017). *The High-Tech Potential.* New York: Routledge.
- Gyourko, J., Mayer, C., & Sinai, T. (2013). Superstar cities. *American Economic Journal: Economic Policy*, 5(4), 167-99.
- Goworowska, J., & Gardner, T. K. (2012). Historical Migration of the Young, Single, and College Educated, 1965 to 2000. *US Census Bureau, Population Division.* Available at: https://www.census.gov/content/dam/Census/library/working-papers/2012/demo/POP-twps0094.pdf [Accessed 20 November 2018]
- Green, L., Myerson, J., Lichtman, D., Rosen, S., & Fry, A. (1996). Temporal discounting in choice between delayed rewards: the role of age and income. *Psychology and aging*, 11(1), 79.
- Greenwood, M. J. (1997). Internal migration in developed countries. *Handbook of population and family economics*, 1, 647-720.
- Greenwood, M. J., Hunt, G. L., & McDowell, J. M. (1986). Migration and employment change: empirical evidence on the spatial and temporal dimensions of the linkage. *Journal of Regional Science*, 26(2), 223-234.
- Greenwood, M. J., Hunt, G. L., Rickman, D. S., & Treyz, G. I. (1991). Migration, regional equilibrium, and the estimation of compensating differentials. *The American Economic Review*, 81(5), 1382-1390.
- Gunderson, R. J., & Sorenson, D. J. (2010). An examination of domestic migration from California counties. *Journal of Regional Analysis & Policy*, 40(1), 34.
- Helderman, A. C., Van Ham, M., & Mulder, C. H. (2006). Migration and home ownership. *Tijdschrift voor Economische en Sociale Geografie*, 97(2), 111-125.
- Hersch, M. H. (2015). Equitable growth and Southern California's aerospace industry. *Washington Center for Equitable Growth*. Available at:

https://equitablegrowth.org/research-paper/equitable-growth-and-southern-californias-aerospace-industry/?longform=true [Accessed 20 December 2018]

- Herzog Jr, H. W., & Schlottmann, A. M. (1986). What can be learned from the recent migrants? *Growth and Change*, 17(1), 37-50.
- Hoehn, J. P., Berger, M. C., & Blomquist, G. C. (1987). A hedonic model of interregional wages, rents, and amenity values. *Journal of Regional Science*, 27(4), 605-620.
- Hsieh, C., and Moretti, E. (2015). Why Do Cities Matter? Local Growth and Aggregate Growth. *Kreisman Working Papers Series in Housing Law and Policy* No. 30.
- International Monetary Fund (2017). World Economic Outlook Database. *WEO Data: October 2017 Edition.* Available at: https://www.imf.org/external/pubs/ft/weo/2017/02/weodata/index.aspx [Accessed 15 November 2018].
- Isserman, A. M., Plane, D. A., & McMillen, D. B. (1982). Internal migration in the United States: An evaluation of federal data. *Review of Public Data Use*, 10(4), 285-311.
- Johnson, H. (2018). California's Future: population. *Public Policy Institute of California.* Available at: https://www.ppic.org/wp-content/uploads/r-118hj2r.pdf [Accessed 21 December 2018]
- Kodrzycki, Y. K. (2001). Migration of recent college graduates: evidence from the national longitudinal survey of youth. *New England Economic Review*, 13-34.
- Kolko, J. (2015). No, Suburbs Aren't All the Same. The Suburbiest Ones Are Growing Fastest. *Citylab*. Available at: https://www.citylab.com/equity/2015/02/no-suburbs-arent-all-the-same-the-suburbiest-ones-are-growing-fastest/385183/ [Accessed 16 January 2019]
- Kolko, J. (2016). Urban revival? Not for most Americans. [Blog] *Terner Center for Housing Innovation: UC Berkeley.* Available at: https://ternercenter.berkeley.edu/blog/urban-revival-not-for-most-americans [Accessed 16 January 2019]
- Kurtz, R. A., & Eicher, J. B. (1958). Fringe and suburb: a confusion of concepts. *Social Forces*, 32-37.
- Lacy, K. (2016). The new sociology of suburbs: A research agenda for analysis of emerging trends. *Annual Review of Sociology*, 42, 369-384.
- Lee, B. A., Oropesa, R. S., & Kanan, J. W. (1994). Neighborhood context and residential mobility. *Demography*, 31(2), 249-270.
- Lewis, K., & Burd-Sharps, S. (2013). Halve the Gap by 2013: Youth Disconnection in America's Cities. *Social Science Research Council.* Available at: http://www.measureofamerica.org/halve-the-gap-2030/ [Accessed 16 January 2019]
- Li, W., & Yao, R. (2007). The life-cycle effects of house price changes. *Journal of Money, Credit and banking*, *39*(6), 1375-1409.
- London School of Economics (2016). *London's wealthiest families feel they are being 'pushed out of elite neighbourhoods'.* Available at: http://www.lse.ac.uk/website-archive/newsAndMedia/newsArchives/2016/08/Londons-wealthiest-familes-feel-they-are-being-pushed-out-of-elite-neighbourhoods.aspx [Accessed 15 January 2019].
- Long, L. (1988). *Migration and residential mobility in the United States*. New York: Russell Sage Foundation.
- Long, S. J., & Freese, J. (2006). *Regression models for categorical dependent variables using Stata.* College Station: StataCorp LP.

- Longhurst, R. (2003). Semi-structured interviews and focus groups. *Key methods in geography*, 117-132.
- Lowry, I. S. (1966). *Migration and metropolitan growth: two analytical models*. San Francisco: Chandler Pub. Co.
- Lucas Jr, R. E. (1988). On the mechanics of economic development. *Journal of monetary economics*, 22(1), 3-42.
- McCann, P. (2013). *Modern urban and regional economics*. Oxford: Oxford University Press.
- Moos, M. (2016). From gentrification to youthification? The increasing importance of young age in delineating high-density living. *Urban Studies*, 53(14), 2903-2920.
- Mehmetoglu, M., & Jakobsen, T. G. (2016). *Applied statistics using Stata: a guide for the social sciences.* Thousand Oaks: Sage.
- Milne, W. J. (1993). Macroeconomic influences on migration. *Regional Studies*, 27(4), 365-373.
- Molloy, R., Smith, C. L., & Wozniak, A. (2011). Internal migration in the United States. *Journal of Economic perspectives*, *25*(3), 173-96.
- Morrison, P. A. (1971). Chronic movers and the future redistribution of population: a longitudinal analysis. *Demography*, 8(2), 171-184.
- Mulder, C. H. (2006). Population and housing: a two-sided relationship. *Demographic Research*, 15, 401-412.
- Mulder, C. H., & Wagner, M. (1998). First-time home-ownership in the family life course: A West German-Dutch comparison. *Urban studies*, 35(4), 687-713.
- National Low Income Housing Coalition (2017). *Out of reach 2017.* Available at: https://nlihc.org/sites/default/files/oor/OOR_2017_0.pdf [Accessed 22 December 2018]
- Nelson, A. C., & Sanchez, T. W. (1997). Exurban and suburban households: a departure from traditional location theory? *Journal of Housing Research*, 249-276.
- Orange County Business Council & Orange County Development Board. (2017). *Workforce Indicators Report.* Available from: https://www.ocbc.org/research/workforce-indicators-report/ [Accessed 16 March 2018].
- Orange County Business Council & Orange County Development Board. (2018). *Workforce Indicators Report.* Available from: https://www.ocbc.org/wp-content/uploads/2018/11/2019-WIR-Reprint-FINAL-Web.pdf [Accessed 17 November 2018].
- Pickford, O. (2016). Barriers to spatial mobility are barriers to social mobility. *The Urbanist*, 4 April. Available at: https://www.theurbanist.org/2016/04/04/why-seattle-needs-inclusionary-zoning-explained-by-the-bid-rent-curve/ [Accessed: 25 December 2018]
- Quigley, J. M., & Raphael, S. (2004). Is housing unaffordable? Why isn't it more affordable? *Journal of Economic Perspectives*, 18(1), 191-214.
- Pissarides, C. A., & Wadsworth, J. (1989). Unemployment and the inter-regional mobility of labour. *The Economic Journal*, 99(397), 739-755.
- Rabionet, S. E. (2011). How I Learned to Design and Conduct Semi-Structured Interviews: An Ongoing and Continuous Journey. *Qualitative Report*, 16(2), 563-566.
- Ratcliffe, M., Burd, C., Holder, K., & Fields, A. (2016). Defining Rural at the U.S. Census Bureau. *U.S. Census Bureau*. Available at:

https://www2.census.gov/geo/pdfs/reference/ua/Defining_Rural.pdf [Accessed 22 December 2018]

- Rhode, P. W. (2001). The evolution of California manufacturing. *Public Policy Institute of California.* Available at: https://www.ppic.org/content/pubs/report/R_1001PRR.pdf [Accessed 20 December 2018]
- Roback, J. (1982). Wages, rents, and the quality of life. *Journal of political Economy*, 90(6), 1257-1278.
- Roback, J. (1988). Wages, rents, and amenities: differences among workers and regions. *Economic Inquiry*, 26(1), 23-41.
- Roseman, C. C. (1977). Changing migration patterns within the United States. *Resource papers for college geography*, (77-2).
- Rosen, S. (1979). Wage-based indexes of urban quality of life. *Current issues in urban economics*, 74-104.
- Ruggles, S. J., Genadek, K., Goeken, R., Grover, J., & Sobek, M. (2017). *Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]*. Minneapolis: University of Minnesota.
- Saks, R. E., & Wozniak, A. (2011). Labor reallocation over the business cycle: New evidence from internal migration. *Journal of Labor Economics*, 29(4), 697-739.
- Sale, J. E., Lohfeld, L. H., & Brazil, K. (2002). Revisiting the quantitative-qualitative debate: Implications for mixed-methods research. *Quality and quantity*, 36(1), 43-53.
- Speare, A., Goldstein, S., & Frey, W. (1975). *Residential Mobility, Migration and Metropolitan Change.* Cambridge, MA: Ballinger.
- Squires, G. D. (2002). Urban sprawl and the uneven development of metropolitan America. *Urban sprawl: Causes, consequences, and policy responses*, 1-22.
- Stark, O., & Taylor, J. E. (1991). Migration incentives, migration types: The role of relative deprivation. *The economic journal*, 101(408), 1163-1178.
- Starr, k. (2007). *California: A history.* New York: Random House
- The Economist (2018). Booming Seattle struggles to stay affordable. *The Economist.* Available at: https://www.economist.com/united-states/2018/08/11/booming-seattle-struggles-to-stay-affordable [Accessed 29 August 2018]
- Thompson, D. (2015). The Richest Cities for Young People: 1980 vs. Today. *The Atlantic.* Available at: https://www.theatlantic.com/business/archive/2015/02/for-great-american-cities-the-rich-dont-always-get-richer/385513/ [Accessed 7 August 2018]
- University of California (2017). *Quick facts about rural California.* Available at: https://ucanr.edu/sites/UC_CCP/files/125967.pdf [Accessed 22 December 2018]
- U.S. Census Bureau (2000). Profile of Selected Economic Characteristics: 2000. *Census 2000 Summary File 3 (SF 3)*. Available at: https://www.census.gov/geo/reference/ua/urban-rural-2010.html [Accessed 01 January 2019]
- U.S. Census Bureau (2010). *2010 Census Urban and Rural Classification and Urban Area Criteria*. Available at: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=CF [Accessed 15 July 2018]
- U.S. Census Bureau (2016). Selected economic characteristics. *2013-2017 American Community Survey 5-Year Estimates.* Available at:

https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=CF [Accessed 15 July 2018]

- U.S. Census Bureau (2016). Tenure by age of householder: Occupied housing units. *American Community Survey 5-Year Estimate.* Available at: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_1YR_B25007&prodType=table [Accessed 9 July 2018]
- U.S. Census Bureau (2017a). Geographical mobility in the past year by age for current residence in the United States. *American Community Survey 1-Year Estimates.* Available at: https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_17_1YR_B07001&prodType=table [Accessed 19 November 2018]
- U.S. Department of Agriculture (2017). 2017 State Agriculture Overview. *National Agriculture Statistics Service Information.* Available at: https://www.nass.usda.gov/Quick_Stats/Ag_Overview/stateOverview.php?state=CALIFORNIA [Accessed 20 December 2018]
- U.S. Department of Commerce (2017). Gross domestic product (GDP) by state. *Bureau of Economic Analysis*. Available at: https://apps.bea.gov/iTable/iTable.cfm?0=1200&isuri=1&reqid=70&step=10&1=1&2=200&3=sic&4=1&5=xx&6=-1&7=-1&8=-1&9=70&10=levels#reqid=70&step=10&isuri=1&7003=200&7035=-1&7004=naics&7005=1&7006=xx&7036=-1&7001=1200&7002=1&7090=70&7007=-1&7093=levels [Accessed 12 December 2018)
- U.S. Department of Commerce (2018). Metropolitan and Micropolitan Statistical Areas of the United States and Puerto Rico August 2017. *Economics and Statistics Administration U.S. Census Bureau*. Available at: https://www2.census.gov/geo/maps/metroarea/us_wall/Aug2017/cbsa_us_0817.pdf [Accessed 8 August 2018].
- U.S. Department of Labor Statistics (2017). *Current employment statistics.* Available at: https://www.labormarketinfo.edd.ca.gov/data/employment-projections.html [Accessed 20 December 2018].
- Valentine, G. (2005) 'Tell me about . . . using interviews as a research methodology', in R. Flowerdew and D. Martin (eds) *Methods in Human Geography: A Guide for Students Doing a Research Project*. London: Pearson Education
- Vara, V. (2015). *How California Bested Texas.* New Yorker, January 8.
- Venhorst, V. A., Edzes, A. J. E., Broersma, L., & Dijk, J. V. (2011). *Brain drain of brain gain?* Hoger opgeleiden in grote steden in Nederland. s.n
- Von Thunen, J. H. (1826). *Der isolierte Staat in Beziehung auf landtschaft under nationalokonomie. Hamburg*.(English translation by CM Wartenberg. Von Thunen's Isolated State. Oxford: Pergamon Press, 1966).
- Whisler, R. L., Waldorf, B. S., Mulligan, G. F., & Plane, D. A. (2008). Quality of life and the migration of the college-educated: a life-course approach. *Growth and Change*, 39(1), 58-94.
- Wilson, W. J. (1987). *The truly disadvantaged: The inner city, the underclass, and public policy*. Chicago: University of Chicago.
- Winchester, H. P. (1999). Interviews and Questionnaires as Mixed Methods in Population Geography: TheCase of Lone Fathers in Newcastle, Australia. *The professional geographer*, 51(1), 60-67.

- Woetzel, J., Mischke, J., Peloquin, S., Weisfield, D. (2016). A tool kit to close California's housing gap: 3.5 million homes by 2025. *McKinsey Global Institute*. Available at: https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Urbanization/Closing%20Californias%20housing%20gap/Closing-Californias-housing-gap-Full-report.ashx [Accessed: 22 December 2018]
- Woetzel, J., Ram, S., Mischke, J., Garemo, N.,& Sankhe, S. (2014). A blueprint for addressing the global affordable housing challenge. *McKinsey Global Institute.* Available at: https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Urbanization/Tackling%20the%20worlds%20affordable%20housing%20challenge/MGI_Affordable_housing_Executive%20summary_October%202014.ashx [Accessed 23 December 2018]
- Wolpert, J. (1965). Behavioral aspects of the decision to migrate. *Papers in Regional Science*, 15(1), 159-169.

Appendix A – List of counties in the analysis

Counties with Housing Affordability Index Rates (California Association of Realtors, 2018a)						Counties with ASEC variables (Flood et al., 2018)		
2000-2018	*2010-2018*	*2014-2018*	*2015-2018*	*2017-2018*	*Missing*	*ASEC*		*Missing*
Alameda	Kings	San Joaquin	Amador	Calaveras	Alpine	Alameda	Sacramento	Alpine
Contra Costa	Madera	Stanislaus	Butte	Lassen	Colusa	Butte	San Bernardino	Amador
Fresno	Monterey		El Dorado	Mono	Del Norte	Contra Costa	San Diego	Calaveras
Los Angeles	Napa		Humboldt	Nevada	Glenn	El Dorado	San Francisco	Colusa
Marin	Placer		Lake	Plumas	Inyo	Fresno	San Joaquin	Del Norte
Merced	Solano		Kern	Tuolumne	Imperial	Humboldt	San Luis Obispo	Glenn
Orange	Tulare		Mariposa		Modoc	Imperial	San Mateo	Inyo
Riverside			Mendocino		Sierra	Kern	Santa Barbara	Lake
Sacramento			San Benito		Trinity	Kings	Santa Clara	Lassen
San Bernardino			Shasta			Los Angeles	Santa Cruz	Mariposa
San Diego			Siskiyou			Madera	Shasta	Mendocino
San Francisco			Sutter			Marin	Solano	Modoc
San Luis Obispo			Tehama			Merced	Sonoma	Mono
San Mateo			Yolo			Monterey	Stanislaus	Nevada
Santa Barbara			Yuba			Napa	Tulare	Plumas
Santa Clara						Orange	Ventura	San Benito
Santa Cruz						Placer	Yolo	Sierra
Sonoma						Riverside		Siskiyou
Ventura								Sutter
								Tehama
								Trinity
								Tuolumne
								Yuba

Source: Survey data from the Annual Social and Economic Supplement (Flood et al., 2018) and Traditional Housing Affordability Index (California Association of Realtors, 2018a).

Counties in the final sample of this research

2000-2018		*2010-2018*	*2014-2018*	*2015-2018*	*Missing*	
Alameda	San Diego	Kings	San Joaquin	El Dorado	Alpine	Modoc
Contra Costa	San Francisco	Madera	Stanislaus	Kern	Amador	Mono
Fresno	San Luis Obispo	Monterey		Yolo	Butte	Nevada
Los Angeles	San Mateo	Napa			Calaveras	Plumas
Marin	Santa Barbara	Placer			Colusa	San Benito
Merced	Santa Clara	Solano			Del Norte	Shasta
Orange	Santa Cruz	Tulare			Glenn	Sierra
Riverside	Sonoma				Inyo	Siskiyou
Sacramento	Ventura				Imperial	Sutter
San Bernardino					Humboldt	Tehama
					Lake	Trinity
					Lassen	Tuolumne
					Mariposa	Yuba
					Mendocino	

Source: Own elaboration on availability of data from ASEC (Flood et al., 2018) and CAR (California Association of Realtors, 2018a).

Appendix B – Housing affordability index rates per county in the analysis

Table B1 Counties with Housing Affordability Index rates for 2000-2018, part 1.

Month/QTR	Alameda	Contra Costa	Fresno	LA	Marin	Merced	Orange	Riverside	Sacramento	San Bernardino	San Diego	San Francisco	San Luis Obispo	San Mateo	Santa Barbara	Santa Clara	Santa Cruz	Sonoma	Ventura
jan-00	24%	15%	53%	39%	16%	54%	31%	45%	57%	52%	27%	16%	28%	15%	15%	24%	21%	24%	35%
feb-00	23%	18%	58%	37%	15%	50%	28%	43%	57%	52%	28%	13%	25%	15%	19%	20%	17%	23%	35%
mrt-00	20%	13%	59%	37%	15%	51%	28%	46%	55%	52%	25%	12%	26%	12%	21%	16%	14%	22%	32%
apr-00	18%	11%	55%	35%	15%	53%	27%	45%	54%	52%	25%	10%	21%	12%	26%	15%	16%	21%	29%
mei-00	18%	12%	53%	36%	14%	54%	26%	42%	53%	51%	22%	10%	26%	11%	19%	16%	13%	18%	32%
jun-00	16%	12%	52%	33%	15%	49%	25%	44%	51%	50%	23%	9%	24%	13%	25%	16%	17%	15%	25%
jul-00	18%	13%	56%	36%	14%	47%	27%	44%	51%	50%	23%	12%	23%	13%	32%	16%	13%	18%	30%
aug-00	20%	12%	51%	33%	15%	47%	26%	43%	51%	51%	23%	10%	20%	13%	23%	18%	14%	18%	31%
sep-00	20%	14%	52%	34%	15%	49%	27%	44%	52%	50%	24%	12%	21%	14%	24%	20%	14%	16%	31%
okt-00	18%	13%	49%	35%	15%	43%	28%	43%	52%	49%	23%	11%	20%	13%	36%	18%	15%	18%	31%
nov-00	17%	14%	55%	34%	15%	50%	28%	42%	52%	52%	23%	10%	25%	15%	14%	17%	15%	17%	33%
dec-00	18%	14%	54%	36%	15%	46%	28%	44%	53%	52%	24%	10%	24%	14%	25%	18%	14%	19%	33%
jan-01	21%	14%	58%	38%	16%	47%	29%	48%	54%	55%	25%	11%	23%	14%	19%	17%	16%	20%	36%
feb-01	20%	15%	58%	38%	16%	44%	31%	45%	52%	56%	25%	11%	22%	13%	25%	19%	15%	16%	38%
mrt-01	19%	13%	55%	37%	17%	48%	29%	45%	51%	52%	26%	11%	#N/A	15%	24%	19%	15%	18%	37%
apr-01	19%	13%	53%	38%	15%	46%	28%	42%	50%	54%	25%	11%	26%	17%	21%	22%	14%	21%	37%
mei-01	22%	14%	53%	37%	17%	48%	29%	41%	49%	54%	26%	12%	28%	17%	26%	22%	15%	19%	33%
jun-01	21%	11%	53%	35%	16%	40%	29%	40%	48%	52%	25%	12%	23%	17%	26%	21%	18%	19%	34%
jul-01	21%	14%	50%	35%	17%	44%	28%	43%	49%	52%	24%	12%	24%	17%	14%	23%	15%	18%	31%
aug-01	22%	14%	52%	34%	17%	41%	27%	44%	49%	52%	24%	13%	27%	18%	17%	25%	19%	21%	35%
sep-01	25%	16%	50%	35%	18%	44%	30%	41%	51%	55%	25%	15%	26%	19%	26%	27%	20%	21%	33%
okt-01	25%	17%	53%	36%	20%	43%	31%	46%	52%	54%	29%	16%	29%	19%	27%	30%	20%	22%	36%
nov-01	26%	18%	52%	36%	20%	40%	33%	45%	52%	54%	29%	14%	28%	21%	28%	29%	20%	23%	35%
dec-01	25%	18%	52%	36%	20%	50%	31%	44%	51%	53%	27%	15%	29%	20%	22%	29%	21%	24%	39%
jan-02	25%	16%	51%	34%	20%	42%	31%	44%	50%	54%	27%	17%	25%	20%	21%	29%	20%	23%	37%
feb-02	21%	15%	50%	33%	18%	40%	26%	42%	50%	55%	25%	14%	22%	17%	20%	23%	17%	19%	36%
mrt-02	21%	11%	51%	33%	16%	40%	25%	39%	48%	53%	23%	12%	20%	16%	22%	22%	16%	19%	36%
apr-02	17%	11%	52%	31%	14%	40%	23%	36%	46%	53%	22%	13%	21%	15%	10%	21%	16%	17%	32%
mei-02	18%	9%	50%	31%	15%	38%	22%	39%	44%	52%	20%	10%	21%	16%	21%	20%	15%	18%	32%
jun-02	17%	10%	46%	31%	16%	37%	22%	37%	44%	51%	20%	12%	20%	16%	21%	20%	16%	17%	33%
jul-02	18%	10%	47%	31%	16%	38%	22%	38%	43%	51%	20%	13%	23%	17%	17%	22%	16%	18%	28%
aug-02	19%	11%	47%	29%	17%	39%	23%	39%	42%	52%	20%	14%	20%	17%	16%	25%	16%	18%	30%
sep-02	20%	12%	47%	31%	18%	35%	24%	41%	42%	50%	20%	14%	24%	18%	22%	28%	17%	20%	30%
okt-02	21%	13%	45%	31%	18%	36%	24%	40%	43%	52%	22%	14%	24%	20%	15%	27%	18%	22%	31%
nov-02	22%	14%	46%	29%	18%	34%	25%	37%	43%	50%	22%	14%	23%	19%	22%	28%	20%	22%	32%
dec-02	22%	12%	48%	31%	21%	36%	25%	37%	43%	51%	22%	15%	22%	21%	21%	28%	20%	22%	30%
jan-03	24%	16%	46%	30%	21%	32%	25%	38%	43%	48%	22%	15%	22%	20%	21%	31%	20%	21%	31%
feb-03	21%	13%	46%	31%	17%	34%	21%	36%	43%	50%	22%	11%	18%	18%	18%	25%	16%	19%	32%
mrt-03	19%	12%	44%	31%	17%	36%	22%	34%	43%	49%	21%	12%	19%	17%	15%	25%	17%	20%	31%
apr-03	20%	10%	44%	29%	17%	33%	21%	33%	43%	49%	22%	11%	17%	18%	12%	26%	16%	18%	23%
mei-03	20%	11%	44%	28%	18%	31%	21%	35%	42%	47%	21%	12%	18%	17%	19%	26%	17%	17%	26%
jun-03	20%	12%	43%	29%	18%	33%	22%	33%	41%	46%	20%	13%	20%	17%	21%	26%	17%	16%	27%
jul-03	20%	13%	42%	27%	18%	33%	21%	33%	42%	45%	19%	12%	19%	18%	12%	28%	19%	17%	27%
aug-03	20%	11%	40%	24%	18%	30%	18%	31%	39%	44%	17%	12%	18%	18%	13%	26%	18%	16%	24%
sep-03	19%	10%	41%	24%	18%	31%	18%	32%	38%	42%	17%	12%	20%	18%	14%	26%	17%	16%	22%
okt-03	20%	12%	41%	24%	19%	27%	18%	29%	38%	39%	16%	12%	18%	18%	19%	26%	18%	17%	25%
nov-03	20%	13%	37%	23%	18%	30%	18%	31%	38%	43%	16%	12%	19%	18%	20%	27%	19%	15%	25%
dec-03	20%	12%	36%	23%	17%	29%	18%	28%	37%	39%	15%	12%	16%	18%	19%	26%	18%	15%	23%
jan-04	21%	14%	42%	23%	19%	31%	20%	27%	39%	41%	16%	12%	17%	18%	15%	29%	18%	16%	24%
feb-04	19%	10%	35%	22%	18%	33%	17%	27%	36%	38%	15%	12%	14%	18%	16%	27%	17%	16%	22%
mrt-04	20%	12%	35%	22%	17%	31%	14%	22%	34%	36%	15%	12%	20%	17%	14%	23%	17%	16%	20%
apr-04	18%	11%	34%	20%	16%	30%	13%	22%	33%	35%	14%	12%	20%	17%	13%	24%	15%	15%	17%
mei-04	15%	10%	31%	18%	14%	25%	11%	20%	30%	32%	11%	11%	16%	15%	7%	22%	14%	14%	15%
jun-04	14%	10%	29%	17%	13%	23%	11%	17%	26%	29%	10%	10%	14%	14%	10%	21%	16%	13%	13%
jul-04	14%	10%	28%	17%	14%	24%	11%	17%	26%	25%	10%	11%	14%	15%	14%	21%	13%	13%	14%
aug-04	15%	10%	27%	17%	14%	21%	11%	18%	26%	28%	10%	11%	14%	15%	14%	23%	14%	13%	14%
sep-04	15%	10%	27%	17%	15%	22%	13%	18%	25%	28%	11%	12%	13%	16%	6%	23%	14%	13%	16%
okt-04	14%	10%	27%	17%	16%	19%	13%	18%	25%	26%	12%	11%	14%	16%	9%	23%	13%	13%	18%
nov-04	15%	10%	28%	17%	14%	20%	13%	18%	24%	27%	12%	11%	14%	15%	7%	22%	15%	12%	17%
dec-04	14%	10%	27%	17%	14%	19%	13%	17%	24%	26%	11%	11%	14%	15%	9%	21%	14%	11%	16%
jan-05	14%	10%	26%	17%	13%	19%	12%	15%	24%	23%	11%	11%	13%	15%	7%	21%	12%	11%	14%
feb-05	13%	10%	26%	17%	13%	17%	11%	17%	24%	24%	11%	10%	13%	14%	8%	20%	12%	10%	15%
mrt-05	12%	10%	24%	17%	11%	16%	11%	19%	23%	23%	10%	9%	13%	11%	9%	19%	12%	9%	14%
apr-05	11%	10%	24%	16%	12%	16%	11%	19%	21%	24%	10%	8%	11%	10%	10%	18%	12%	8%	13%
mei-05	11%	10%	20%	15%	12%	14%	11%	18%	20%	24%	9%	8%	11%	11%	6%	18%	11%	7%	13%
jun-05	11%	10%	20%	15%	10%	11%	11%	14%	21%	20%	9%	9%	10%	12%	7%	18%	10%	7%	13%
jul-05	11%	10%	21%	14%	11%	11%	11%	15%	20%	19%	9%	9%	10%	12%	7%	18%	11%	7%	13%
aug-05	11%	10%	18%	12%	11%	10%	11%	18%	19%	22%	9%	9%	10%	12%	6%	18%	10%	7%	13%
sep-05	11%	10%	19%	13%	13%	9%	11%	18%	20%	22%	9%	9%	8%	13%	9%	19%	11%	7%	13%
okt-05	11%	10%	18%	12%	12%	9%	11%	18%	19%	22%	9%	9%	8%	12%	9%	18%	11%	7%	13%
nov-05	11%	10%	17%	11%	12%	10%	11%	17%	19%	21%	8%	9%	7%	12%	7%	18%	10%	7%	13%
dec-05	11%	10%	19%	12%	10%	10%	10%	16%	19%	20%	9%	9%	10%	13%	6%	18%	11%	7%	13%

Source: Traditional Housing Affordability Index, California Association of Realtors (2018a)

Table B2 Counties with Housing Affordability Index rates for 2000-2018, part 2.

Month/QTR	Alameda	Contra Costa	Fresno	LA	Marin	Merced	Orange	Riverside	Sacramento	San Bernardino	San Diego	San Francisco	San Luis Obispo	San Mateo	Santa Barbara	Santa Clara	Santa Cruz	Sonoma	Ventura
2006,1	12%	11%	25%	11%	12%	20%	11%	17%	22%	22%	9%	11%	8%	12%	7%	15%	10%	9%	11%
2006,2	23%	15%	39%	9%	20%	36%	10%	32%	21%	38%	9%	16%	7%	18%	6%	13%	19%	22%	10%
2006,3	23%	17%	40%	9%	21%	41%	11%	25%	21%	28%	9%	17%	7%	20%	7%	14%	18%	23%	10%
2006,4	25%	23%	42%	9%	22%	41%	11%	34%	24%	37%	11%	19%	11%	20%	8%	15%	21%	25%	11%
2007,1	12%	10%	28%	10%	11%	29%	12%	19%	25%	24%	11%	10%	10%	11%	6%	14%	9%	11%	12%
2007,2	23%	16%	44%	10%	19%	46%	11%	35%	26%	41%	10%	15%	7%	18%	6%	11%	18%	24%	11%
2007,3	11%	8%	28%	9%	10%	31%	11%	19%	28%	27%	10%	8%	9%	8%	6%	11%	9%	10%	11%
2007,4	12%	12%	31%	15%	11%	34%	13%	24%	35%	28%	16%	11%	14%	10%	9%	13%	10%	19%	13%
2008,1	17%	15%	39%	20%	14%	42%	18%	37%	46%	41%	24%	12%	16%	12%	17%	16%	13%	23%	25%
2008,2	20%	14%	43%	22%	11%	53%	20%	41%	51%	46%	25%	12%	18%	13%	23%	17%	13%	26%	29%
2008,3	21%	15%	48%	25%	14%	60%	25%	49%	54%	52%	29%	13%	20%	15%	25%	20%	16%	31%	30%
2008,4	31%	17%	57%	30%	19%	69%	31%	57%	61%	56%	39%	17%	27%	20%	37%	32%	27%	38%	40%
2009,1	45%	32%	65%	41%	24%	75%	37%	65%	69%	68%	43%	24%	31%	32%	47%	44%	36%	45%	47%
2009,2	38%	25%	66%	41%	23%	76%	35%	68%	68%	73%	41%	21%	32%	21%	38%	40%	29%	44%	41%
2009,3	32%	22%	63%	35%	22%	73%	32%	65%	65%	70%	36%	19%	27%	21%	33%	33%	26%	39%	35%
2009,4	33%	24%	63%	35%	24%	74%	33%	65%	66%	69%	38%	19%	29%	21%	31%	32%	26%	39%	39%
2010,1	34%	25%	64%	43%	24%	77%	28%	63%	68%	75%	38%	21%	34%	22%	34%	34%	27%	40%	39%
2010,2	30%	21%	62%	43%	21%	73%	27%	60%	65%	68%	37%	19%	31%	20%	25%	28%	26%	40%	39%
2010,3	31%	21%	65%	38%	23%	75%	28%	61%	68%	74%	38%	22%	36%	21%	26%	30%	28%	40%	40%
2010,4	33%	26%	67%	43%	25%	76%	33%	64%	70%	76%	40%	22%	37%	25%	32%	35%	29%	44%	41%
2011,1	35%	30%	69%	46%	27%	76%	33%	63%	71%	76%	40%	25%	40%	30%	37%	38%	34%	47%	44%
2011,2	35%	26%	70%	46%	24%	76%	31%	64%	72%	77%	41%	24%	37%	22%	35%	32%	32%	46%	41%
2011,3	36%	27%	69%	42%	25%	74%	33%	65%	72%	77%	42%	26%	40%	29%	38%	34%	32%	46%	45%
2011,4	39%	37%	71%	48%	29%	77%	38%	66%	74%	78%	46%	26%	41%	29%	41%	40%	37%	51%	49%
2012,1	45%	36%	72%	51%	32%	77%	39%	66%	74%	78%	46%	29%	41%	33%	42%	42%	35%	51%	50%
2012,2	38%	26%	71%	49%	27%	77%	35%	65%	74%	78%	44%	24%	41%	23%	31%	32%	34%	49%	48%
2012,3	34%	28%	69%	42%	27%	74%	34%	63%	73%	77%	43%	25%	37%	24%	30%	32%	30%	46%	47%
2012,4	36%	31%	70%	44%	28%	74%	34%	62%	71%	76%	43%	22%	40%	24%	27%	32%	34%	46%	48%
2013,1	32%	28%	64%	42%	24%	68%	28%	54%	62%	72%	38%	23%	34%	23%	24%	30%	29%	36%	42%
2013,2	25%	20%	61%	37%	20%	65%	23%	49%	56%	69%	32%	17%	30%	17%	18%	24%	20%	29%	36%
2013,3	21%	22%	56%	27%	18%	60%	20%	45%	50%	64%	27%	16%	23%	15%	18%	22%	18%	31%	30%
2013,4	23%	24%	55%	30%	19%	60%	20%	43%	51%	62%	28%	16%	24%	16%	18%	23%	17%	30%	31%
2014,1	22%	24%	54%	31%	15%	58%	21%	42%	50%	61%	27%	15%	24%	14%	18%	22%	20%	29%	29%
2014,2	18%	19%	53%	30%	14%	57%	20%	41%	48%	58%	26%	10%	23%	14%	18%	19%	18%	29%	28%
2014,3	21%	20%	53%	25%	15%	55%	20%	41%	48%	57%	25%	12%	24%	15%	14%	21%	17%	29%	27%
2014,4	20%	23%	53%	28%	15%	53%	21%	41%	49%	57%	27%	11%	26%	15%	21%	22%	17%	29%	29%
2015,1	25%	41%	50%	31%	19%	59%	22%	42%	49%	58%	27%	12%	29%	14%	16%	21%	21%	30%	28%
2015,2	20%	33%	49%	29%	17%	54%	20%	40%	47%	55%	25%	10%	27%	13%	14%	18%	20%	26%	25%
2015,3	20%	34%	49%	24%	19%	56%	20%	40%	47%	54%	24%	11%	27%	13%	17%	19%	19%	26%	25%
2015,4	22%	38%	49%	26%	17%	55%	21%	39%	46%	53%	25%	11%	26%	14%	20%	20%	21%	27%	26%
2016,1	25%	40%	52%	31%	20%	55%	23%	42%	48%	56%	29%	13%	27%	16%	20%	22%	18%	28%	30%
2016,2	22%	34%	47%	29%	17%	50%	21%	40%	45%	55%	27%	13%	26%	14%	18%	19%	16%	25%	32%
2016,3	23%	36%	48%	25%	18%	49%	21%	40%	43%	54%	25%	13%	23%	14%	19%	21%	17%	26%	29%
2016,4	23%	38%	49%	27%	19%	47%	21%	40%	45%	53%	26%	13%	26%	14%	21%	21%	16%	25%	30%
2017.1	21%	37%	48%	29%	18%	50%	21%	39%	46%	52%	28%	13%	26%	15%	14%	19%	17%	25%	28%
2017,2	19%	31%	47%	28%	17%	48%	21%	39%	45%	51%	26%	12%	26%	14%	16%	17%	17%	25%	27%
2017,3	20%	33%	45%	22%	18%	47%	21%	38%	43%	51%	26%	13%	23%	15%	20%	17%	17%	25%	27%
2017,4	20%	34%	46%	25%	18%	48%	21%	38%	43%	50%	26%	12%	24%	14%	18%	15%	17%	23%	26%
2018,1	22%	36%	49%	28%	18%	43%	21%	39%	44%	52%	26%	15%	25%	15%	22%	17%	15%	21%	31%
2018,2	16%	29%	47%	26%	18%	42%	20%	37%	41%	49%	23%	14%	22%	14%	20%	16%	12%	20%	28%
2018,3	18%	32%	46%	22%	19%	40%	20%	37%	42%	48%	23%	15%	21%	14%	26%	17%	12%	22%	28%

Source: Traditional Housing Affordability Index, California Association of Realtors (2018a)

Table B13 Average Housing Affordability Index per year for counties with data for 2000-2018.

Year	Alameda	Contra Costa	Fresno	LA	Marin	Merced	Orange	Riverside	Sacramento	San Bernardino	San Diego	San Francisco	San Luis Obispo	San Mateo	Santa Barbara	Santa Clara	Santa Cruz	Sonoma	Ventura
2000	19%	13%	54%	35%	15%	49%	27%	44%	53%	51%	24%	11%	24%	13%	23%	18%	15%	19%	31%
2001	22%	15%	53%	36%	17%	45%	30%	44%	51%	54%	26%	13%	26%	17%	23%	24%	17%	20%	35%
2002	20%	12%	48%	31%	17%	38%	24%	39%	45%	52%	22%	14%	22%	18%	19%	24%	17%	20%	32%
2003	20%	12%	42%	27%	18%	32%	20%	33%	41%	45%	19%	12%	19%	18%	17%	27%	18%	17%	26%
2004	16%	11%	31%	19%	15%	25%	13%	20%	29%	31%	12%	11%	15%	16%	11%	23%	15%	14%	17%
2005	12%	10%	21%	14%	12%	13%	11%	17%	21%	22%	9%	9%	10%	12%	8%	19%	11%	8%	13%
2006	21%	17%	37%	10%	19%	35%	11%	27%	22%	31%	10%	16%	8%	18%	7%	14%	17%	20%	11%
2007	15%	12%	33%	11%	13%	35%	12%	24%	29%	30%	12%	11%	10%	12%	7%	12%	12%	16%	12%
2008	22%	15%	47%	24%	15%	56%	24%	46%	53%	49%	29%	14%	20%	15%	26%	21%	17%	30%	31%
2009	37%	26%	64%	38%	23%	75%	34%	66%	67%	70%	40%	21%	30%	24%	37%	37%	29%	42%	41%
2010	32%	23%	65%	42%	23%	75%	29%	62%	68%	73%	38%	21%	35%	22%	29%	32%	28%	41%	40%
2011	36%	30%	70%	46%	26%	76%	34%	65%	72%	77%	42%	25%	40%	28%	38%	36%	34%	48%	45%
2012	38%	30%	71%	46%	28%	76%	36%	64%	73%	77%	44%	25%	40%	26%	33%	35%	33%	48%	48%
2013	25%	24%	59%	34%	20%	63%	23%	48%	55%	67%	31%	18%	28%	18%	20%	25%	21%	32%	35%
2014	20%	22%	54%	29%	15%	56%	20%	41%	49%	58%	26%	12%	24%	15%	18%	21%	18%	29%	28%
2015	22%	37%	49%	28%	18%	56%	21%	40%	47%	55%	25%	11%	27%	13%	17%	20%	20%	28%	26%
2016	23%	37%	49%	28%	18%	50%	22%	41%	45%	55%	27%	13%	25%	15%	19%	20%	17%	26%	30%
2017	20%	34%	47%	26%	18%	48%	21%	39%	44%	51%	26%	13%	25%	15%	17%	17%	17%	24%	27%
2018	19%	32%	47%	25%	18%	42%	20%	38%	43%	50%	24%	15%	23%	14%	23%	17%	13%	21%	29%

Source: Own calculation of Traditional Housing Affordability Index, California Association of Realtors (2018a)

Table B4 Counties with Housing Affordability Index rates for 2010-2018.

Month/QT	Kings	Madera	Monterey	Napa	Placer	Solano	Tulare
2010,1	64%	65%	56%	44%	58%	70%	66%
2010,2	62%	67%	52%	42%	58%	69%	66%
2010,3	64%	68%	56%	41%	60%	71%	68%
2010,4	66%	70%	58%	47%	62%	71%	71%
2011,1	71%	68%	57%	48%	64%	74%	72%
2011,2	72%	72%	56%	47%	64%	75%	73%
2011,3	76%	74%	56%	48%	65%	75%	73%
2011,4	75%	75%	56%	50%	67%	76%	73%
2012,1	75%	77%	54%	50%	67%	77%	75%
2012,2	74%	74%	55%	50%	65%	77%	73%
2012,3	74%	76%	52%	45%	64%	77%	73%
2012,4	76%	74%	50%	47%	63%	73%	71%
2013,1	70%	77%	44%	39%	57%	67%	70%
2013,2	70%	71%	35%	34%	51%	63%	66%
2013,3	62%	63%	30%	28%	46%	56%	61%
2013,4	63%	67%	29%	29%	47%	56%	60%
2014,1	64%	61%	23%	28%	45%	53%	59%
2014,2	64%	50%	26%	25%	44%	50%	56%
2014,3	64%	51%	27%	21%	44%	49%	56%
2014,4	64%	50%	27%	24%	45%	50%	56%
2015,1	62%	50%	28%	29%	46%	48%	56%
2015,2	62%	49%	27%	23%	44%	46%	53%
2015,3	60%	49%	27%	23%	44%	45%	53%
2015,4	61%	48%	25%	21%	44%	46%	54%
2016,1	58%	50%	26%	23%	48%	47%	52%
2016,2	52%	50%	24%	24%	46%	44%	49%
2016,3	56%	46%	23%	24%	45%	43%	48%
2016,4	55%	48%	26%	25%	46%	44%	48%
2017.1	53%	47%	23%	24%	45%	45%	52%
2017,2	52%	44%	21%	25%	43%	44%	52%
2017,3	52%	44%	22%	26%	44%	43%	51%
2017,4	52%	49%	21%	25%	44%	44%	52%
2018,1	52%	50%	23%	28%	44%	42%	50%
2018,2	50%	52%	19%	25%	41%	38%	48%
2018,3	51%	48%	20%	24%	42%	38%	47%

Source: Traditional Housing Affordability Index, California Association of Realtors (2018a)

Table B5 Average Housing Affordability Index per year for counties with data for 2010-2018.

Year	Kings	Madera	Monterey	Napa	Placer	Solano	Tulare
2010	64%	68%	56%	44%	60%	70%	68%
2011	74%	72%	56%	48%	65%	75%	73%
2012	75%	75%	53%	48%	65%	76%	73%
2013	66%	70%	35%	33%	50%	61%	64%
2014	64%	53%	26%	25%	45%	50%	57%
2015	61%	49%	27%	24%	44%	46%	54%
2016	55%	48%	25%	24%	46%	45%	49%
2017	52%	46%	22%	25%	44%	44%	52%
2018	51%	50%	21%	26%	42%	39%	48%

Source: Own calculation of Traditional Housing Affordability Index, California Association of Realtors (2018a)

Table B6 Counties with Housing Affordability Index rates for 2014-2018 & 2015-2018.

Month/QTR	Butte	El Dorado	Humboldt	Kern	San Joaquin	Shasta	Stanislaus	Yolo
2014,1					42%		46%	
2014,2					40%		45%	
2014,3					40%		44%	
2014,4					41%		44%	
2015,1	46%	45%	43%	47%	48%	53%	48%	42%
2015,2	43%	41%	43%	54%	45%	51%	47%	37%
2015,3	43%	43%	42%	53%	45%	51%	46%	37%
2015,4	43%	44%	42%	55%	46%	51%	45%	38%
2016,1	43%	39%	42%	55%	47%	49%	50%	37%
2016,2	42%	37%	41%	54%	44%	47%	47%	35%
2016,3	41%	39%	38%	54%	43%	46%	47%	35%
2016,4	40%	38%	39%	54%	43%	45%	47%	36%
2017.1	41%	43%	36%	55%	45%	49%	48%	37%
2017,2	39%	40%	36%	54%	43%	47%	47%	35%
2017,3	40%	41%	33%	53%	40%	48%	46%	34%
2017,4	39%	42%	34%	54%	41%	47%	45%	34%
2018,1	41%	42%	36%	56%	40%	49%	48%	41%
2018,2	38%	38%	33%	53%	38%	46%	45%	33%
2018,3	39%	41%	33%	53%	38%	44%	45%	35%

Source: Traditional Housing Affordability Index, California Association of Realtors (2018a)

Table B7 Average Housing Affordability Index per year for counties with data for 2014-2018 & 2015-2018

Year	Butte	El Dorado	Humboldt	Kern	San Joaquin	Shasta	Stanislaus	Yolo
2014					41%		45%	
2015	44%	43%	42%	52%	46%	51%	47%	39%
2016	42%	38%	40%	54%	44%	47%	48%	36%
2017	40%	42%	35%	54%	42%	48%	46%	35%
2018	39%	40%	34%	54%	38%	46%	46%	37%

Source: Own calculation of Traditional Housing Affordability Index, California Association of Realtors (2018a)

Appendix C – Methodology Traditional Housing Affordability Index

Source: Directly copied from California Association of Realtors, 2018b

THE ASSUMPTIONS AND METHODOLOGY USED TO CALCULATE C.A.R.'S TRADITIONAL HOUSING AFFORDABILITY INDEX (HAI)

Step 1. MEDIAN PRICE: C.A.R.'s housing affordability index is based on the median price of existing single-family homes sold from C.A.R.'s monthly existing home sales survey. Starting in 1987, this survey is based on reports of closed escrow sales from 80 Boards or more of REALTORS® and multiple listing services around the state. Prior to 1987, the survey was based on reports from 45 Boards.

Step 2. DOWNPAYMENT: It is assumed that a household can make a 20 percent downpayment on the median-priced home. Therefore, the loan amount needed to purchase a home would be 80 percent of the median home sales price.

Step 3. INTEREST RATE: Using the national average effective mortgage interest rate on all fixed and adjustable rate mortgages. This is represented by the effective composite rate for previously occupied homes, which is reported monthly by the Federal Housing Finance Board.

Step 4.The monthly payment for PRINCIPAL, INTEREST, TAXES AND INSURANCE (PITI) is computed as the sum of three parts:

-Monthly mortgage payment, based on the terms of the mortgage in Steps 2 & 3.
-Monthly PROPERTY TAXES are assumed to be 1 percent of the median home sales price divided by 12.
-Monthly INSURANCE PAYMENTS on the house are assumed to be 0.38 percent of the median home sales price divided by 12.

The results of these three calculations are added together to find the PITI or total monthly payment for a household that buys the median priced home.

Step 5. It is then assumed that the monthly PITI can be no more than 30 percent of a household's income. Thus, the monthly housing payment is divided by .3 to come up with the MINIMUM INCOME NEEDED TO QUALIFY FOR A LOAN on the median-priced home.

Step 6. Starting in 1988, data for the distribution of households by various income ranges was obtained from Claritas. INCOME DISTRIBUTION figures were developed based on the projected percent change in the annual median household income. Prior to 1988, household income utilized in the housing affordability index was based on projections by C.A.R. using the 1980 census data as a base.

Step 7. The minimum income amount calculated in Step 5 is multiplied by 12 to determine the minimum annual income needed to qualify. This amount is compared to the income distribution of households. The percent of the households with incomes greater than or equal to the minimum income becomes the HOUSING AFFORDABILITY INDEX (HAI).

NOTE: The quarterly HAI series begins in 2006, prior to that the series was monthly. The quarterly HAI for a given geographic area in a particular quarter is based upon the quarterly median price for that area as well as the quarterly income distribution for that area.

Appendix D – Syntax

```
* IPUMS - CPS Dataset* --> Includes both the ASEC and BMS

*ONLY KEEP ASEC SURVEY*
*Drop all observations that do not have migration values, so only keep
the ASEC survey from the CPS in the database (also see thesis "Surveys
in the dataset")
keep if asecflag==1

*Migration variables*

gen migration=.
*stands for migration (inter or intra)
replace migration=0 if (migrate1==1)
replace migration=1 if (migrate1==3)
replace migration=2 if (migrate1==4) | (migrate1==5)
label variable migration "Internal_migration"
label define migration_lbl 0"Same_house" 1"Intracounty_migration"
2"Intercounty_migration"
label values migration migration_lbl

gen intermig=.
*stands for migration to another county (=internal migration);
migration status based upon migrate1. If someone has migrated the
previous year to another county =1; no migration =0
replace intermig = 0 if ( migrate1==1)
*These persons have stayed in the same house
replace intermig = 0 if (migrate1==3)
*These persons have stayed in the same county (= not seen as internal
migration, but only as moving)
replace intermig = 1 if (migrate1==4)
*Moved within state, different county
replace intermig = 1 if (migrate1==5)
*Moved between states
label variable intermig "migration to another county"
label define intermig_lbl 0"No" 1"Yes"
label values intermig intermig_lbl

gen intramig=.
*stands for migration within the county; migration status based upon
migrate1. If someone has migrated the previous year within the same
county =1; other =0
replace intramig = 0 if ( migrate1==1)
replace intramig = 1 if (migrate1==3)
replace intramig = 0 if (migrate1==4)
replace intramig = 0 if (migrate1==5)
label variable intramig "migration within county"
label define intramig_lbl 0"No" 1"Yes"
label values intramig intramig_lbl

*Counties
label define county_lbl 6001"Alameda County" 6007"Butte" 6013"Contra
Costa" 6017"El Dorado" 6019"Fresno" 6023"Humboldt" 6029"Kern"
6031"Kings" 6037"Los Angeles" 6039"Madera" 6041"Marin" 6047"Merced"
6053"Monterey" 6055"Napa" 6059"Orange" 6061"Placer" 6065"Riverside"
6067"Sacramento" 6071"San Bernardino" 6073"San Diego" 6075"San
Francisco" 6077"San Joaquin" 6079"San Luis Obispo" 6081"San Mateo"
```

```
6083"Santa Barbara" 6085"Santa Clara" 6087"Santa Cruz" 6089"Shasta"
6095"Solano" 6097"Sonoma" 6099"Stanislaus" 6107"Tulare" 6111"Ventura"
6113"Yolo"
label values county county_lbl

*Drop rural counties
drop if county == 6007
drop if county == 6023
drop if county == 6089

***IMPORTING HOUSING AFFORDABILITY INDEX***
generate HAI =.
label variable HAI "Housing affordability index"
*Alameda County (6001)
replace HAI = 19 if (county ==6001) & (year == 2000)
replace HAI = 22 if (county ==6001) & (year == 2001)
replace HAI = 20 if (county ==6001) & (year == 2002)
replace HAI = 20 if (county ==6001) & (year == 2003)
replace HAI = 16 if (county ==6001) & (year == 2004)
replace HAI = 12 if (county ==6001) & (year == 2005)
replace HAI = 21 if (county ==6001) & (year == 2006)
replace HAI = 15 if (county ==6001) & (year == 2007)
replace HAI = 22 if (county ==6001) & (year == 2008)
replace HAI = 37 if (county ==6001) & (year == 2009)
replace HAI = 32 if (county ==6001) & (year == 2010)
replace HAI = 36 if (county ==6001) & (year == 2011)
replace HAI = 38 if (county ==6001) & (year == 2012)
replace HAI = 25 if (county ==6001) & (year == 2013)
replace HAI = 20 if (county ==6001) & (year == 2014)
replace HAI = 22 if (county ==6001) & (year == 2015)
replace HAI = 23 if (county ==6001) & (year == 2016)
replace HAI = 20 if (county ==6001) & (year == 2017)
replace HAI = 19 if (county ==6001) & (year == 2018)
*Butte
replace HAI = 44 if (county ==6007) & (year == 2015)
replace HAI = 42 if (county ==6007) & (year == 2016)
replace HAI = 40 if (county ==6007) & (year == 2017)
replace HAI = 39 if (county ==6007) & (year == 2018)
* Contra Costa
replace HAI = 13 if (county ==6013) & (year == 2000)
replace HAI = 15 if (county ==6013) & (year == 2001)
replace HAI = 12 if (county ==6013) & (year == 2002)
replace HAI = 12 if (county ==6013) & (year == 2003)
replace HAI = 11 if (county ==6013) & (year == 2004)
replace HAI = 10 if (county ==6013) & (year == 2005)
replace HAI = 17 if (county ==6013) & (year == 2006)
replace HAI = 12 if (county ==6013) & (year == 2007)
replace HAI = 15 if (county ==6013) & (year == 2008)
replace HAI = 26 if (county ==6013) & (year == 2009)
replace HAI = 23 if (county ==6013) & (year == 2010)
replace HAI = 30 if (county ==6013) & (year == 2011)
replace HAI = 30 if (county ==6013) & (year == 2012)
replace HAI = 24 if (county ==6013) & (year == 2013)
replace HAI = 22 if (county ==6013) & (year == 2014)
replace HAI = 37 if (county ==6013) & (year == 2015)
replace HAI = 37 if (county ==6013) & (year == 2016)
replace HAI = 34 if (county ==6013) & (year == 2017)
replace HAI = 32 if (county ==6013) & (year == 2018)
```

```
*El Dorado 017
replace HAI = 43 if (county ==6017) & (year == 2015)
replace HAI = 38 if (county ==6017) & (year == 2016)
replace HAI = 42 if (county ==6017) & (year == 2017)
replace HAI = 40 if (county ==6017) & (year == 2018)
*Fresno    019
replace HAI = 54 if (county ==6019) & (year == 2000)
replace HAI = 53 if (county ==6019) & (year == 2001)
replace HAI = 48 if (county ==6019) & (year == 2002)
replace HAI = 42 if (county ==6019) & (year == 2003)
replace HAI = 31 if (county ==6019) & (year == 2004)
replace HAI = 21 if (county ==6019) & (year == 2005)
replace HAI = 37 if (county ==6019) & (year == 2006)
replace HAI = 33 if (county ==6019) & (year == 2007)
replace HAI = 47 if (county ==6019) & (year == 2008)
replace HAI = 64 if (county ==6019) & (year == 2009)
replace HAI = 65 if (county ==6019) & (year == 2010)
replace HAI = 70 if (county ==6019) & (year == 2011)
replace HAI = 71 if (county ==6019) & (year == 2012)
replace HAI = 59 if (county ==6019) & (year == 2013)
replace HAI = 54 if (county ==6019) & (year == 2014)
replace HAI = 49 if (county ==6019) & (year == 2015)
replace HAI = 49 if (county ==6019) & (year == 2016)
replace HAI = 47 if (county ==6019) & (year == 2017)
replace HAI = 47 if (county ==6019) & (year == 2018)
*Humboldt  023
replace HAI = 42 if (county ==6023) & (year == 2015)
replace HAI = 40 if (county ==6023) & (year == 2016)
replace HAI = 35 if (county ==6023) & (year == 2017)
replace HAI = 34 if (county ==6023) & (year == 2018)
*Kern 029
replace HAI = 52 if (county ==6029) & (year == 2015)
replace HAI = 54 if (county ==6029) & (year == 2016)
replace HAI = 54 if (county ==6029) & (year == 2017)
replace HAI = 54 if (county ==6029) & (year == 2018)
*Kings     031
replace HAI = 64 if (county ==6031) & (year == 2010)
replace HAI = 74 if (county ==6031) & (year == 2011)
replace HAI = 75 if (county ==6031) & (year == 2012)
replace HAI = 66 if (county ==6031) & (year == 2013)
replace HAI = 64 if (county ==6031) & (year == 2014)
replace HAI = 61 if (county ==6031) & (year == 2015)
replace HAI = 55 if (county ==6031) & (year == 2016)
replace HAI = 52 if (county ==6031) & (year == 2017)
replace HAI = 51 if (county ==6031) & (year == 2018)
*Los Angeles     037
replace HAI = 35 if (county ==6037) & (year == 2000)
replace HAI = 36 if (county ==6037) & (year == 2001)
replace HAI = 31 if (county ==6037) & (year == 2002)
replace HAI = 27 if (county ==6037) & (year == 2003)
replace HAI = 19 if (county ==6037) & (year == 2004)
replace HAI = 14 if (county ==6037) & (year == 2005)
replace HAI = 10 if (county ==6037) & (year == 2006)
replace HAI = 11 if (county ==6037) & (year == 2007)
replace HAI = 24 if (county ==6037) & (year == 2008)
replace HAI = 38 if (county ==6037) & (year == 2009)
replace HAI = 42 if (county ==6037) & (year == 2010)
replace HAI = 46 if (county ==6037) & (year == 2011)
```

```
replace HAI = 46 if (county ==6037) & (year == 2012)
replace HAI = 34 if (county ==6037) & (year == 2013)
replace HAI = 29 if (county ==6037) & (year == 2014)
replace HAI = 28 if (county ==6037) & (year == 2015)
replace HAI = 28 if (county ==6037) & (year == 2016)
replace HAI = 26 if (county ==6037) & (year == 2017)
replace HAI = 25 if (county ==6037) & (year == 2018)
*Madera    039
replace HAI = 68 if (county ==6039) & (year == 2010)
replace HAI = 72 if (county ==6039) & (year == 2011)
replace HAI = 75 if (county ==6039) & (year == 2012)
replace HAI = 70 if (county ==6039) & (year == 2013)
replace HAI = 53 if (county ==6039) & (year == 2014)
replace HAI = 49 if (county ==6039) & (year == 2015)
replace HAI = 48 if (county ==6039) & (year == 2016)
replace HAI = 46 if (county ==6039) & (year == 2017)
replace HAI = 50 if (county ==6039) & (year == 2018)
*Marin     041
replace HAI = 15 if (county ==6041) & (year == 2000)
replace HAI = 17 if (county ==6041) & (year == 2001)
replace HAI = 17 if (county ==6041) & (year == 2002)
replace HAI = 18 if (county ==6041) & (year == 2003)
replace HAI = 15 if (county ==6041) & (year == 2004)
replace HAI = 12 if (county ==6041) & (year == 2005)
replace HAI = 19 if (county ==6041) & (year == 2006)
replace HAI = 13 if (county ==6041) & (year == 2007)
replace HAI = 15 if (county ==6041) & (year == 2008)
replace HAI = 23 if (county ==6041) & (year == 2009)
replace HAI = 23 if (county ==6041) & (year == 2010)
replace HAI = 26 if (county ==6041) & (year == 2011)
replace HAI = 28 if (county ==6041) & (year == 2012)
replace HAI = 20 if (county ==6041) & (year == 2013)
replace HAI = 15 if (county ==6041) & (year == 2014)
replace HAI = 18 if (county ==6041) & (year == 2015)
replace HAI = 18 if (county ==6041) & (year == 2016)
replace HAI = 18 if (county ==6041) & (year == 2017)
replace HAI = 18 if (county ==6041) & (year == 2018)
*Merced    047
replace HAI = 49 if (county ==6047) & (year == 2000)
replace HAI = 45 if (county ==6047) & (year == 2001)
replace HAI = 38 if (county ==6047) & (year == 2002)
replace HAI = 32 if (county ==6047) & (year == 2003)
replace HAI = 25 if (county ==6047) & (year == 2004)
replace HAI = 13 if (county ==6047) & (year == 2005)
replace HAI = 35 if (county ==6047) & (year == 2006)
replace HAI = 35 if (county ==6047) & (year == 2007)
replace HAI = 56 if (county ==6047) & (year == 2008)
replace HAI = 75 if (county ==6047) & (year == 2009)
replace HAI = 75 if (county ==6047) & (year == 2010)
replace HAI = 76 if (county ==6047) & (year == 2011)
replace HAI = 76 if (county ==6047) & (year == 2012)
replace HAI = 63 if (county ==6047) & (year == 2013)
replace HAI = 56 if (county ==6047) & (year == 2014)
replace HAI = 56 if (county ==6047) & (year == 2015)
replace HAI = 50 if (county ==6047) & (year == 2016)
replace HAI = 48 if (county ==6047) & (year == 2017)
replace HAI = 42 if (county ==6047) & (year == 2018)
*Monterey  053
```

```
replace HAI = 56 if (county ==6053) & (year == 2010)
replace HAI = 56 if (county ==6053) & (year == 2011)
replace HAI = 53 if (county ==6053) & (year == 2012)
replace HAI = 35 if (county ==6053) & (year == 2013)
replace HAI = 26 if (county ==6053) & (year == 2014)
replace HAI = 27 if (county ==6053) & (year == 2015)
replace HAI = 25 if (county ==6053) & (year == 2016)
replace HAI = 22 if (county ==6053) & (year == 2017)
replace HAI = 21 if (county ==6053) & (year == 2018)
*Napa 055
replace HAI = 44 if (county ==6055) & (year == 2010)
replace HAI = 48 if (county ==6055) & (year == 2011)
replace HAI = 48 if (county ==6055) & (year == 2012)
replace HAI = 33 if (county ==6055) & (year == 2013)
replace HAI = 25 if (county ==6055) & (year == 2014)
replace HAI = 24 if (county ==6055) & (year == 2015)
replace HAI = 24 if (county ==6055) & (year == 2016)
replace HAI = 25 if (county ==6055) & (year == 2017)
replace HAI = 26 if (county ==6055) & (year == 2018)
*Orange    059
replace HAI = 27 if (county ==6059) & (year == 2000)
replace HAI = 30 if (county ==6059) & (year == 2001)
replace HAI = 24 if (county ==6059) & (year == 2002)
replace HAI = 20 if (county ==6059) & (year == 2003)
replace HAI = 13 if (county ==6059) & (year == 2004)
replace HAI = 11 if (county ==6059) & (year == 2005)
replace HAI = 11 if (county ==6059) & (year == 2006)
replace HAI = 12 if (county ==6059) & (year == 2007)
replace HAI = 24 if (county ==6059) & (year == 2008)
replace HAI = 34 if (county ==6059) & (year == 2009)
replace HAI = 29 if (county ==6059) & (year == 2010)
replace HAI = 34 if (county ==6059) & (year == 2011)
replace HAI = 36 if (county ==6059) & (year == 2012)
replace HAI = 23 if (county ==6059) & (year == 2013)
replace HAI = 20 if (county ==6059) & (year == 2014)
replace HAI = 21 if (county ==6059) & (year == 2015)
replace HAI = 22 if (county ==6059) & (year == 2016)
replace HAI = 21 if (county ==6059) & (year == 2017)
replace HAI = 20 if (county ==6059) & (year == 2018)
*Placer    061
replace HAI = 60 if (county ==6061) & (year == 2010)
replace HAI = 65 if (county ==6061) & (year == 2011)
replace HAI = 65 if (county ==6061) & (year == 2012)
replace HAI = 50 if (county ==6061) & (year == 2013)
replace HAI = 45 if (county ==6061) & (year == 2014)
replace HAI = 44 if (county ==6061) & (year == 2015)
replace HAI = 46 if (county ==6061) & (year == 2016)
replace HAI = 44 if (county ==6061) & (year == 2017)
replace HAI = 42 if (county ==6061) & (year == 2018)
*Riverside 065
replace HAI = 44 if (county ==6065) & (year == 2000)
replace HAI = 44 if (county ==6065) & (year == 2001)
replace HAI = 39 if (county ==6065) & (year == 2002)
replace HAI = 33 if (county ==6065) & (year == 2003)
replace HAI = 20 if (county ==6065) & (year == 2004)
replace HAI = 17 if (county ==6065) & (year == 2005)
replace HAI = 27 if (county ==6065) & (year == 2006)
replace HAI = 24 if (county ==6065) & (year == 2007)
```

```
replace HAI = 46 if (county ==6065) & (year == 2008)
replace HAI = 66 if (county ==6065) & (year == 2009)
replace HAI = 62 if (county ==6065) & (year == 2010)
replace HAI = 65 if (county ==6065) & (year == 2011)
replace HAI = 64 if (county ==6065) & (year == 2012)
replace HAI = 48 if (county ==6065) & (year == 2013)
replace HAI = 41 if (county ==6065) & (year == 2014)
replace HAI = 40 if (county ==6065) & (year == 2015)
replace HAI = 41 if (county ==6065) & (year == 2016)
replace HAI = 39 if (county ==6065) & (year == 2017)
replace HAI = 38 if (county ==6065) & (year == 2018)
*Sacramento 067
replace HAI = 53 if (county ==6067) & (year == 2000)
replace HAI = 51 if (county ==6067) & (year == 2001)
replace HAI = 45 if (county ==6067) & (year == 2002)
replace HAI = 41 if (county ==6067) & (year == 2003)
replace HAI = 29 if (county ==6067) & (year == 2004)
replace HAI = 21 if (county ==6067) & (year == 2005)
replace HAI = 22 if (county ==6067) & (year == 2006)
replace HAI = 29 if (county ==6067) & (year == 2007)
replace HAI = 53 if (county ==6067) & (year == 2008)
replace HAI = 67 if (county ==6067) & (year == 2009)
replace HAI = 68 if (county ==6067) & (year == 2010)
replace HAI = 72 if (county ==6067) & (year == 2011)
replace HAI = 73 if (county ==6067) & (year == 2012)
replace HAI = 55 if (county ==6067) & (year == 2013)
replace HAI = 49 if (county ==6067) & (year == 2014)
replace HAI = 47 if (county ==6067) & (year == 2015)
replace HAI = 45 if (county ==6067) & (year == 2016)
replace HAI = 44 if (county ==6067) & (year == 2017)
replace HAI = 43 if (county ==6067) & (year == 2018)
*San Bernardino  071
replace HAI = 51 if (county ==6071) & (year == 2000)
replace HAI = 54 if (county ==6071) & (year == 2001)
replace HAI = 52 if (county ==6071) & (year == 2002)
replace HAI = 45 if (county ==6071) & (year == 2003)
replace HAI = 31 if (county ==6071) & (year == 2004)
replace HAI = 22 if (county ==6071) & (year == 2005)
replace HAI = 31 if (county ==6071) & (year == 2006)
replace HAI = 30 if (county ==6071) & (year == 2007)
replace HAI = 49 if (county ==6071) & (year == 2008)
replace HAI = 70 if (county ==6071) & (year == 2009)
replace HAI = 73 if (county ==6071) & (year == 2010)
replace HAI = 77 if (county ==6071) & (year == 2011)
replace HAI = 77 if (county ==6071) & (year == 2012)
replace HAI = 67 if (county ==6071) & (year == 2013)
replace HAI = 58 if (county ==6071) & (year == 2014)
replace HAI = 55 if (county ==6071) & (year == 2015)
replace HAI = 55 if (county ==6071) & (year == 2016)
replace HAI = 51 if (county ==6071) & (year == 2017)
replace HAI = 50 if (county ==6071) & (year == 2018)
*San Diego 073
replace HAI = 24 if (county ==6073) & (year == 2000)
replace HAI = 26 if (county ==6073) & (year == 2001)
replace HAI = 22 if (county ==6073) & (year == 2002)
replace HAI = 19 if (county ==6073) & (year == 2003)
replace HAI = 12 if (county ==6073) & (year == 2004)
replace HAI = 9 if (county ==6073) & (year == 2005)
```

```
replace HAI = 10 if (county ==6073) & (year == 2006)
replace HAI = 12 if (county ==6073) & (year == 2007)
replace HAI = 29 if (county ==6073) & (year == 2008)
replace HAI = 40 if (county ==6073) & (year == 2009)
replace HAI = 38 if (county ==6073) & (year == 2010)
replace HAI = 42 if (county ==6073) & (year == 2011)
replace HAI = 44 if (county ==6073) & (year == 2012)
replace HAI = 31 if (county ==6073) & (year == 2013)
replace HAI = 26 if (county ==6073) & (year == 2014)
replace HAI = 25 if (county ==6073) & (year == 2015)
replace HAI = 27 if (county ==6073) & (year == 2016)
replace HAI = 26 if (county ==6073) & (year == 2017)
replace HAI = 24 if (county ==6073) & (year == 2018)
*San Francisco   075
replace HAI = 11 if (county ==6075) & (year == 2000)
replace HAI = 13 if (county ==6075) & (year == 2001)
replace HAI = 14 if (county ==6075) & (year == 2002)
replace HAI = 12 if (county ==6075) & (year == 2003)
replace HAI = 11 if (county ==6075) & (year == 2004)
replace HAI = 9 if (county ==6075) & (year == 2005)
replace HAI = 16 if (county ==6075) & (year == 2006)
replace HAI = 11 if (county ==6075) & (year == 2007)
replace HAI = 14 if (county ==6075) & (year == 2008)
replace HAI = 21 if (county ==6075) & (year == 2009)
replace HAI = 21 if (county ==6075) & (year == 2010)
replace HAI = 25 if (county ==6075) & (year == 2011)
replace HAI = 25 if (county ==6075) & (year == 2012)
replace HAI = 18 if (county ==6075) & (year == 2013)
replace HAI = 12 if (county ==6075) & (year == 2014)
replace HAI = 11 if (county ==6075) & (year == 2015)
replace HAI = 13 if (county ==6075) & (year == 2016)
replace HAI = 12 if (county ==6075) & (year == 2017)
replace HAI = 15 if (county ==6075) & (year == 2018)
*San Joaquin     077
replace HAI = 41 if (county ==6077) & (year == 2014)
replace HAI = 46 if (county ==6077) & (year == 2015)
replace HAI = 44 if (county ==6077) & (year == 2016)
replace HAI = 42 if (county ==6077) & (year == 2017)
replace HAI = 38 if (county ==6077) & (year == 2018)
*San Luis Obispo 079
replace HAI = 24 if (county ==6079) & (year == 2000)
replace HAI = 26 if (county ==6079) & (year == 2001)
replace HAI = 22 if (county ==6079) & (year == 2002)
replace HAI = 19 if (county ==6079) & (year == 2003)
replace HAI = 15 if (county ==6079) & (year == 2004)
replace HAI = 10 if (county ==6079) & (year == 2005)
replace HAI = 8 if (county ==6079) & (year == 2006)
replace HAI = 10 if (county ==6079) & (year == 2007)
replace HAI = 20 if (county ==6079) & (year == 2008)
replace HAI = 30 if (county ==6079) & (year == 2009)
replace HAI = 35 if (county ==6079) & (year == 2010)
replace HAI = 40 if (county ==6079) & (year == 2011)
replace HAI = 40 if (county ==6079) & (year == 2012)
replace HAI = 28 if (county ==6079) & (year == 2013)
replace HAI = 24 if (county ==6079) & (year == 2014)
replace HAI = 27 if (county ==6079) & (year == 2015)
replace HAI = 25 if (county ==6079) & (year == 2016)
replace HAI = 25 if (county ==6079) & (year == 2017)
```

```
replace HAI = 23 if (county ==6079) & (year == 2018)
*San Mateo 081
replace HAI = 13 if (county ==6081) & (year == 2000)
replace HAI = 17 if (county ==6081) & (year == 2001)
replace HAI = 18 if (county ==6081) & (year == 2002)
replace HAI = 18 if (county ==6081) & (year == 2003)
replace HAI = 16 if (county ==6081) & (year == 2004)
replace HAI = 12 if (county ==6081) & (year == 2005)
replace HAI = 18 if (county ==6081) & (year == 2006)
replace HAI = 12 if (county ==6081) & (year == 2007)
replace HAI = 15 if (county ==6081) & (year == 2008)
replace HAI = 24 if (county ==6081) & (year == 2009)
replace HAI = 22 if (county ==6081) & (year == 2010)
replace HAI = 28 if (county ==6081) & (year == 2011)
replace HAI = 26 if (county ==6081) & (year == 2012)
replace HAI = 18 if (county ==6081) & (year == 2013)
replace HAI = 15 if (county ==6081) & (year == 2014)
replace HAI = 13 if (county ==6081) & (year == 2015)
replace HAI = 15 if (county ==6081) & (year == 2016)
replace HAI = 15 if (county ==6081) & (year == 2017)
replace HAI = 14 if (county ==6081) & (year == 2018)
*Santa Barbara   083
replace HAI = 23 if (county ==6083) & (year == 2000)
replace HAI = 23 if (county ==6083) & (year == 2001)
replace HAI = 19 if (county ==6083) & (year == 2002)
replace HAI = 17 if (county ==6083) & (year == 2003)
replace HAI = 11 if (county ==6083) & (year == 2004)
replace HAI = 8 if (county ==6083) & (year == 2005)
replace HAI = 7 if (county ==6083) & (year == 2006)
replace HAI = 7 if (county ==6083) & (year == 2007)
replace HAI = 26 if (county ==6083) & (year == 2008)
replace HAI = 37 if (county ==6083) & (year == 2009)
replace HAI = 29 if (county ==6083) & (year == 2010)
replace HAI = 38 if (county ==6083) & (year == 2011)
replace HAI = 33 if (county ==6083) & (year == 2012)
replace HAI = 20 if (county ==6083) & (year == 2013)
replace HAI = 18 if (county ==6083) & (year == 2014)
replace HAI = 17 if (county ==6083) & (year == 2015)
replace HAI = 19 if (county ==6083) & (year == 2016)
replace HAI = 17 if (county ==6083) & (year == 2017)
replace HAI = 23 if (county ==6083) & (year == 2018)
*Santa Clara    085
replace HAI = 18 if (county ==6085) & (year == 2000)
replace HAI = 24 if (county ==6085) & (year == 2001)
replace HAI = 24 if (county ==6085) & (year == 2002)
replace HAI = 27 if (county ==6085) & (year == 2003)
replace HAI = 23 if (county ==6085) & (year == 2004)
replace HAI = 19 if (county ==6085) & (year == 2005)
replace HAI = 14 if (county ==6085) & (year == 2006)
replace HAI = 12 if (county ==6085) & (year == 2007)
replace HAI = 21 if (county ==6085) & (year == 2008)
replace HAI = 37 if (county ==6085) & (year == 2009)
replace HAI = 32 if (county ==6085) & (year == 2010)
replace HAI = 36 if (county ==6085) & (year == 2011)
replace HAI = 35 if (county ==6085) & (year == 2012)
replace HAI = 25 if (county ==6085) & (year == 2013)
replace HAI = 21 if (county ==6085) & (year == 2014)
replace HAI = 20 if (county ==6085) & (year == 2015)
```

```
replace HAI = 20 if (county ==6085) & (year == 2016)
replace HAI = 17 if (county ==6085) & (year == 2017)
replace HAI = 17 if (county ==6085) & (year == 2018)
*Santa Cruz 087
replace HAI = 15 if (county ==6087) & (year == 2000)
replace HAI = 17 if (county ==6087) & (year == 2001)
replace HAI = 17 if (county ==6087) & (year == 2002)
replace HAI = 18 if (county ==6087) & (year == 2003)
replace HAI = 15 if (county ==6087) & (year == 2004)
replace HAI = 11 if (county ==6087) & (year == 2005)
replace HAI = 17 if (county ==6087) & (year == 2006)
replace HAI = 12 if (county ==6087) & (year == 2007)
replace HAI = 17 if (county ==6087) & (year == 2008)
replace HAI = 29 if (county ==6087) & (year == 2009)
replace HAI = 28 if (county ==6087) & (year == 2010)
replace HAI = 34 if (county ==6087) & (year == 2011)
replace HAI = 33 if (county ==6087) & (year == 2012)
replace HAI = 21 if (county ==6087) & (year == 2013)
replace HAI = 18 if (county ==6087) & (year == 2014)
replace HAI = 20 if (county ==6087) & (year == 2015)
replace HAI = 17 if (county ==6087) & (year == 2016)
replace HAI = 17 if (county ==6087) & (year == 2017)
replace HAI = 13 if (county ==6087) & (year == 2018)
*Shasta    089
replace HAI = 51 if (county ==6089) & (year == 2015)
replace HAI = 47 if (county ==6089) & (year == 2016)
replace HAI = 48 if (county ==6089) & (year == 2017)
replace HAI = 46 if (county ==6089) & (year == 2018)
*Solano    095
replace HAI = 70 if (county ==6095) & (year == 2010)
replace HAI = 75 if (county ==6095) & (year == 2011)
replace HAI = 76 if (county ==6095) & (year == 2012)
replace HAI = 61 if (county ==6095) & (year == 2013)
replace HAI = 50 if (county ==6095) & (year == 2014)
replace HAI = 46 if (county ==6095) & (year == 2015)
replace HAI = 45 if (county ==6095) & (year == 2016)
replace HAI = 44 if (county ==6095) & (year == 2017)
replace HAI = 39 if (county ==6095) & (year == 2018)
*Sonoma    097
replace HAI = 19 if (county ==6097) & (year == 2000)
replace HAI = 20 if (county ==6097) & (year == 2001)
replace HAI = 20 if (county ==6097) & (year == 2002)
replace HAI = 17 if (county ==6097) & (year == 2003)
replace HAI = 14 if (county ==6097) & (year == 2004)
replace HAI = 8 if (county ==6097) & (year == 2005)
replace HAI = 20 if (county ==6097) & (year == 2006)
replace HAI = 16 if (county ==6097) & (year == 2007)
replace HAI = 30 if (county ==6097) & (year == 2008)
replace HAI = 42 if (county ==6097) & (year == 2009)
replace HAI = 41 if (county ==6097) & (year == 2010)
replace HAI = 48 if (county ==6097) & (year == 2011)
replace HAI = 48 if (county ==6097) & (year == 2012)
replace HAI = 32 if (county ==6097) & (year == 2013)
replace HAI = 29 if (county ==6097) & (year == 2014)
replace HAI = 28 if (county ==6097) & (year == 2015)
replace HAI = 26 if (county ==6097) & (year == 2016)
replace HAI = 24 if (county ==6097) & (year == 2017)
replace HAI = 21 if (county ==6097) & (year == 2018)
```

```
*Stanislaus 099
replace HAI = 45 if (county ==6099) & (year == 2014)
replace HAI = 47 if (county ==6099) & (year == 2015)
replace HAI = 48 if (county ==6099) & (year == 2016)
replace HAI = 46 if (county ==6099) & (year == 2017)
replace HAI = 46 if (county ==6099) & (year == 2018)
*Tulare    107
replace HAI = 68 if (county ==6107) & (year == 2010)
replace HAI = 73 if (county ==6107) & (year == 2011)
replace HAI = 73 if (county ==6107) & (year == 2012)
replace HAI = 64 if (county ==6107) & (year == 2013)
replace HAI = 57 if (county ==6107) & (year == 2014)
replace HAI = 54 if (county ==6107) & (year == 2015)
replace HAI = 49 if (county ==6107) & (year == 2016)
replace HAI = 52 if (county ==6107) & (year == 2017)
replace HAI = 48 if (county ==6107) & (year == 2018)
*Ventura   111
replace HAI = 31 if (county ==6111) & (year == 2000)
replace HAI = 35 if (county ==6111) & (year == 2001)
replace HAI = 32 if (county ==6111) & (year == 2002)
replace HAI = 26 if (county ==6111) & (year == 2003)
replace HAI = 17 if (county ==6111) & (year == 2004)
replace HAI = 13 if (county ==6111) & (year == 2005)
replace HAI = 11 if (county ==6111) & (year == 2006)
replace HAI = 12 if (county ==6111) & (year == 2007)
replace HAI = 31 if (county ==6111) & (year == 2008)
replace HAI = 41 if (county ==6111) & (year == 2009)
replace HAI = 40 if (county ==6111) & (year == 2010)
replace HAI = 45 if (county ==6111) & (year == 2011)
replace HAI = 48 if (county ==6111) & (year == 2012)
replace HAI = 35 if (county ==6111) & (year == 2013)
replace HAI = 28 if (county ==6111) & (year == 2014)
replace HAI = 26 if (county ==6111) & (year == 2015)
replace HAI = 30 if (county ==6111) & (year == 2016)
replace HAI = 27 if (county ==6111) & (year == 2017)
replace HAI = 29 if (county ==6111) & (year == 2018)
*Yolo 113
replace HAI = 39 if (county ==6113) & (year == 2015)
replace HAI = 36 if (county ==6113) & (year == 2016)
replace HAI = 35 if (county ==6113) & (year == 2017)
replace HAI = 37 if (county ==6113) & (year == 2018)

*Whymove contains niu value that reflect the persons who did not move -
-> therefore, create new variable without NIU value
gen whymove_new=.
replace whymove_new = 1 if (whymove==01)
replace whymove_new = 2 if (whymove==02)
replace whymove_new = 3 if (whymove==03)
replace whymove_new = 4 if (whymove==04)
replace whymove_new = 5 if (whymove==05)
replace whymove_new = 6 if (whymove==06)
replace whymove_new = 7 if (whymove==07)
replace whymove_new = 8 if (whymove==08)
replace whymove_new = 9 if (whymove==09)
replace whymove_new = 10 if (whymove==10)
replace whymove_new = 11 if (whymove==11)
replace whymove_new = 12 if (whymove==12)
replace whymove_new = 13 if (whymove==13)
```

```
replace whymove_new = 14 if (whymove==14)
replace whymove_new = 15 if (whymove==15)
replace whymove_new = 16 if (whymove==16)
replace whymove_new = 17 if (whymove==17)
replace whymove_new = 18 if (whymove==18)
replace whymove_new = 19 if (whymove==19)
label variable whymove_new "reason for moving"
label define whymove_new_lbl 1"change in marital status" 2"to establish
own household" 3"other family reason" 4"new job or job transfer" 5"to
look for work or lost job" 6"for easier commute" 7"retired" 8"other
job-related reason" 9"wanted to own home, not rent" 10"wanted new or
better housing" 11"wanted better neighborhood" 12"for cheaper housing"
13"other housing reason" 14"attend/leave college" 15"change of climate"
16"health reasons" 17"other reasons" 18"natural disaster"
19"foreclosure or eviction"
label values whymove_new whymove_new_lbl

generate whymove_h=.
replace whymove_h=1 if (whymove==12)
replace whymove_h=0 if (whymove_h==.)
label variable whymove_h "Reason for moving"
label define whymove_h_lbl 0"Other" 1"Looking for cheaper housing"

label values whymove_h whymove_h_lbl

* Creating age cohorts
generate agecohort=.
replace agecohort = 1 if (age==0) | (age==1) | (age==2) | (age==3) |
(age==4)
replace agecohort = 2 if (age==5) | (age==6) | (age==7) |(age==8) |
(age==9)
replace agecohort = 3 if (age==10) | (age==11) | (age==12) |(age==13) |
(age==14)
replace agecohort = 4 if (age==15) | (age==16) | (age==17) |(age==18) |
(age==19)
replace agecohort = 5 if (age==20) | (age==21) | (age==22) |(age==23) |
(age==24)
replace agecohort = 6 if (age==25) | (age==26) | (age==27) |(age==28) |
(age==29)
replace agecohort = 7 if (age==30) | (age==31) | (age==32) |(age==33) |
(age==34)
replace agecohort = 8 if (age==35) | (age==36) | (age==37) |(age==38) |
(age==39)
replace agecohort = 9 if (age==40) | (age==41) | (age==42) |(age==43) |
(age==44)
replace agecohort = 10 if (age==45) | (age==46) | (age==47) |(age==48)
| (age==49)
replace agecohort = 11 if (age==50) | (age==51) | (age==52) |(age==53)
| (age==54)
replace agecohort = 12 if (age==55) | (age==56) | (age==57) |(age==58)
| (age==59)
replace agecohort = 13 if (age==60) | (age==61) | (age==62) |(age==63)
| (age==64)
replace agecohort = 14 if (age==65) | (age==66) | (age==67) |(age==68)
| (age==69)
replace agecohort = 15 if (age==70) | (age==71) | (age==72) |(age==73)
| (age==74)
replace agecohort = 16 if (age>=75)
```

```
label variable agecohort "age cohorts"
label define agecohort_lbl 1"0-4" 2"5-9" 3"10-14" 4"15-19" 5"20-24"
6"25-29" 7"30-34" 8"35-39" 9"40-44" 10"45-49" 11"50-54" 12"55-59"
13"60-64" 14"65-69" 15"70-74" 16"75+"
label values agecohort agecohort_lbl

*Creating year groups
gen year_3 =.
replace year_3=1 if (year==2000) | (year==2001) |(year==2002)
|(year==2003) |(year==2004) | (year==2005) | (year==2006)
replace year_3=2 if (year==2007) |(year==2008) |(year==2009) |
(year==2010) | (year==2011) |(year==2012)
replace year_3=3 if (year==2013) |(year==2014) | (year==2015) |
(year==2016) |(year==2017) |(year==2018)
label variable year_3 "years"
label define year_3_lbl 1"2000-2006" 2"2007-2012" 3"2013-2018"
label values year_3 year_3_lbl

*Modify hhincome + summary (=tabstat)
xtile hhincome_3 = hhincome, nq(3)
label variable hhincome_3 "Household income"
label define hhincome_3_lbl 1"Lowest quintile" 2"Central quintile"
3"Highest quintile"
label values hhincome_3 hhincome_3_lbl

*Modify bpl
generate bpl_new1=.
replace bpl_new1=1 if (bpl==09900)
replace bpl_new1=2 if (bpl_new1==.)
label variable bpl_new1 "Country of birth"
label define bpl_new1_lbl 1"USA" 2"Other"
label values bpl_new1 bpl_new1_lbl

*Modify empstat
generate empstat_new3=.
replace empstat_new3=1 if (empstat==21) |(empstat==22)
replace empstat_new3=2 if (empstat==01) | (empstat==10) | (empstat==12)
replace empstat_new3=3 if (empstat==30) | (empstat==31) | (empstat==32)
| (empstat==33) | (empstat==34) | (empstat==35) | (empstat==36)
label variable empstat_new3 "employment status"
label define empstat_new3_lbl 1"unemployed" 2"employed" 3"not in the
labor force"
label values empstat_new3 empstat_new3_lbl

*Modify educ
generate educ_3=.
label variable educ_3 "Highest degree completed"
replace educ_3=1 if
(educ==002)|(educ==010)|(educ==020)|(educ==030)|(educ==040)|(educ==050)
|(educ==060)|(educ==071)|(educ==073)
replace educ_3=2 if (educ==091) |(educ==092)| (educ==081)
replace educ_3=3 if (educ==111)|(educ==123)|(educ==124)| (educ==125)
label define educ_3_lbl 1"Less educated" 2"Middle educated/middle
skilled" 3"Higher educated"
label values educ_3 educ_3_lbl

*Modify ownershp
generate ownershp_new=.
```

```
replace ownershp_new=1 if(ownershp==10)
replace ownershp_new=2 if(ownershp==21) | (ownershp==22)
label variable ownershp_new "Tenure"
label define ownershp_new_lbl 1"Owner" 2"Renter"
label values ownershp_new ownershp_new_lbl

*Modify race
generate race_new=.
replace race_new=1 if (race==100)
replace race_new=2 if (race==200)
replace race_new=3 if (race_new==.)
label variable race_new "Race"
label define race_new_lbl 1"White" 2"Black/negro" 3"other"
label values race_new race_new_lbl

*Modify nchild
generate children=.
replace children=0 if (nchild==0)
replace children=1 if (children==.)
label variable children "Presence of children in household"
label define children_lbl 0"None" 1"At least one"
label values children children_lbl

***DATA ANALYSIS***
tab migration whymove_h if agecohort==4, cell
tab migration whymove_h if agecohort==5, cell
tab migration whymove_h if agecohort==6, cell
tab migration whymove_h if agecohort==7, cell
tab migration whymove_h if agecohort==8, cell
tab migration whymove_h if agecohort==9, cell
tab migration whymove_h if agecohort==10, cell
tab migration whymove_h if agecohort==11, cell
tab migration whymove_h if agecohort==12, cell
tab migration whymove_h if agecohort==13, cell
tab migration whymove_h if agecohort==14, cell
tab migration whymove_h if agecohort==15, cell
tab migration whymove_h if agecohort==16, cell

* Descriptive statistics
tab migration if state==06
tabstat HAI if migration==0 if state==06, stat(mean SD)
tabstat HAI if migration==1 if state==06, stat(mean SD)
tabstat HAI if migration==2 if state==06, stat(mean SD)
tab2 year_3 migration if state==06
tab2 agecohort migration if state==06
tab2 sex migration if state==06
tab2 bpl_new1 migration if state==06
tab2 race_new migration if state==06
tab2 marst migration if state==06
tab2 children migration if state==06
tab2 workly migration if state==06
tab2 hhincome_3 migration if state==06
tab2 empstat_new3 migration if state==06
tab2 educ_3 migration if state==06
tab2 ownershp_new migration if state==06

*Correlation matrix
```

```
pwcorr migration HAI workly hhincome_5 empstat_new2 educ_new2 bpl_new1
ownershp_new rentsub pubhous
pwcorr HAI eldorado year_3 sex workly hhincome_3 empstat_new3 educ_4
bpl_new1 race_new ownershp_new marst children county if age_young==1

***Regressions***
mlogit migration HAI i.year_3 i.agecohort i.sex i.bpl_new1 i.race_new
i.marst i.children i.county, vce(cluster county) rrr base(0)
mlogit migration HAI i.year_3 i.agecohort i.sex i.bpl_new1 i.race_new
i.marst i.children i.workly i.hhincome_3 i.empstat_new3 i.educ_3
i.ownershp_new i.county, vce(cluster county) rrr base(0)
*Models for agegroups
mlogit migration HAI i.year_3 i.sex i.workly i.hhincome_3
i.empstat_new3 i.educ_3 i.bpl_new1 i.race_new i.ownershp_new i.marst
i.children i.county if agecohort==4 | agecohort==5, vce(cluster county)
rrr base(0)
mlogit migration HAI i.year_3 i.sex i.workly i.hhincome_3
i.empstat_new3 i.educ_3 i.bpl_new1 i.race_new i.ownershp_new i.marst
i.children i.county if agecohort==6 | agecohort==7, vce(cluster county)
rrr base(0)
mlogit migration HAI i.year_3 i.sex i.workly i.hhincome_3
i.empstat_new3 i.educ_3 i.bpl_new1 i.race_new i.ownershp_new i.marst
i.children i.county if agecohort==8 | agecohort==9, vce(cluster county)
rrr base(0)
mlogit migration HAI i.year_3 i.sex i.workly i.hhincome_3
i.empstat_new3 i.educ_3 i.bpl_new1 i.race_new i.ownershp_new i.marst
i.children i.county if agecohort==10 | agecohort==11 | agecohort==12 |
agecohort==13 | agecohort==14 | agecohort==15 | agecohort==16 ,
vce(cluster county) rrr base(0)
*Models for education:
mlogit migration HAI i.year_3 i.agecohort i.sex i.workly i.hhincome_3
i.empstat_new3 i.bpl_new1 i.race_new i.ownershp_new i.marst i.children
i.county if educ_3==1, vce(cluster county) rrr base(0)
mlogit migration HAI i.year_3 i.agecohort i.sex i.workly i.hhincome_3
i.empstat_new3 i.bpl_new1 i.race_new i.ownershp_new i.marst i.children
i.county if educ_3==2, vce(cluster county) rrr base(0)
mlogit migration HAI i.year_3 i.agecohort i.sex i.workly i.hhincome_3
i.empstat_new3 i.bpl_new1 i.race_new i.ownershp_new i.marst i.children
i.county if educ_3==3, vce(cluster county) rrr base(0)


*In tables:
quietly mlogit migration HAI i.year_3 i.agecohort i.sex i.bpl_new1
i.race_new i.marst i.children i.county, vce(cluster county) rrr base(0)
eststo reg1
quietly mlogit migration HAI i.year_3 i.agecohort i.sex i.bpl_new1
i.race_new i.marst i.children i.workly i.hhincome_3 i.empstat_new3
i.educ_3  i.ownershp_new i.county, vce(cluster county) rrr base(0)
eststo reg2
quietly mlogit migration HAI i.year_3 i.sex i.workly i.hhincome_3
i.empstat_new3 i.educ_3 i.bpl_new1 i.race_new i.ownershp_new i.marst
i.children i.county if agecohort==4 | agecohort==5, vce(cluster county)
rrr base(0)
eststo reg3
quietly mlogit migration HAI i.year_3 i.sex i.workly i.hhincome_3
i.empstat_new3 i.educ_3 i.bpl_new1 i.race_new i.ownershp_new i.marst
i.children i.county if agecohort==6 | agecohort==7, vce(cluster county)
rrr base(0)
```

```
eststo reg4
quietly mlogit migration HAI i.year_3 i.sex i.workly i.hhincome_3
i.empstat_new3 i.educ_3 i.bpl_new1 i.race_new i.ownershp_new i.marst
i.children i.county if agecohort==8 | agecohort==9, vce(cluster county)
rrr base(0)
eststo reg5
quietly mlogit migration HAI i.year_3 i.sex i.workly i.hhincome_3
i.empstat_new3 i.educ_3 i.bpl_new1 i.race_new i.ownershp_new i.marst
i.children i.county if agecohort==10 | agecohort==11 | agecohort==12 |
agecohort==13 | agecohort==14 | agecohort==15 | agecohort==16 ,
vce(cluster county) rrr base(0)
eststo reg6
quietly mlogit migration HAI i.year_3 i.agecohort i.sex i.workly
i.hhincome_3 i.empstat_new3  i.bpl_new1 i.race_new i.ownershp_new
i.marst i.children i.county if educ_3==1, vce(cluster county) rrr
base(0)
eststo reg7
quietly mlogit migration HAI i.year_3 i.agecohort i.sex i.workly
i.hhincome_3 i.empstat_new3 i.bpl_new1 i.race_new i.ownershp_new
i.marst i.children i.county if educ_3==2, vce(cluster county) rrr
base(0)
eststo reg8
quietly mlogit migration HAI i.year_3 i.agecohort i.sex i.workly
i.hhincome_3 i.empstat_new3 i.bpl_new1 i.race_new i.ownershp_new
i.marst i.children i.county if educ_3==3, vce(cluster county) rrr
base(0)
eststo reg9
esttab reg1 reg2 reg3 reg4 reg5 reg6 reg7 reg8 reg9, eform se label
keep ($covariates) star(* 0.10 ** 0.05 *** 0.01 ) title("Logit
reasons") addnote("Source: Flood et al., 2018; California Association
of Realtors, 2018")
```

Appendix E – Interview guide

Note: Even though semi-constructed interviews are usually conducted with a predetermined list of questions, these unfold in a conversational manner offering participants the possibility to explore issues they feel are important. Therefore this document should be used as an assisting guide for conducting the semi-constructed interviews.

Introduction • Introduce the research • Explain relevant theories • Show and explain results from the quantitative analyses **Questions** *Results/analysis* • How do you think about the methodology used for the quantitative analyses? • How would you expect the relationship between housing affordability and internal migration into California counties? • What are your thoughts about the results from the quantitative models? *Housing affordability in California* • The paper of Bean, Brown & Pullés (2018) argues that unaffordable housing is likely to be the driving force behind out-migration patterns from California. How would you describe housing affordability in California? • How important is affordable housing for California's future economic prosperity? • How do you think housing affordability works on internal migration? Is it only the cost of housing or could it also be other factors? *The cause* • What causes the unaffordability of housing in California counties? • Why is it the coastal counties that are so expensive while the inland counties are reasonable affordable? *The affected ones* • Which demographic groups are mainly affected by regional housing affordability? • For age: Which age groups are important to keep or attract in California counties? And why are these important? • For education: How important is it to keep or attract either the higher educated or less educated? And why are these important?	Notes:

The balance • How could California regulate affordable housing in order to keep or attract those groups that are important for a county? • Can there be any negative consequences related to providing affordable housing? (E.g. traffic congestion or pollution?) • How do you think a proper balance could be achieved between providing affordable housing and the potential negative consequences resulting from that?	