Norwegian University
of Life Sciences

Master's Thesis 2020 30 ECTS
School of Economics and Business

Westside, best side?

- A study of converging housing prices in the city of Oslo

Kjetil Hodder Hovden
Master of Science in Economics

Acknowledgements

This thesis marks the end of my time studying a master's degree in economics at the Norwegian University of Life Sciences (NMBU). The years spent at NMBU have been a great experience for me, and I will have many fond memories of this chapter of my life.

As for this master thesis, I would like to thank my supervisor, affiliated Ph.D. candidate Mari Olsen Mamre, for her guidance during this semester. Your knowledge, feedback and suggested literature have been of great help and undoubtedly increased the quality of this thesis.

I would also like to thank Eiendomsverdi AS for sharing their data. Finally, I want thank those who have helped me with proofreading the paper.

Kjetil Hodder Hovden
Ås, December 15, 2020

Abstract

The goal of this thesis was to analyse the housing price pattern in Oslo, and to find evidence of price segmentation between the 15 districts in Oslo. To form the price index used to test for convergence, the weighted repeated sales model created by Case and Shiller (1987) was used. The analysis was based on quarterly data from 1998 to 2019 was used in this analysis. A panel model developed by Phillips and Sul was used to run the convergence tests. This was done for both total prices and prices per square meter.

No evidence for overall convergence was found. When testing for total housing prices there were three convergence clubs and one divergence club. There were also identified three convergence clubs when using prices per square meter. Alna was excluded from any club "membership" when using the latter price estimation.

Apart from one exception, all the convergence clubs have kept their position relative to the other clubs in the time period. For the clubs formed by using housing prices per square meter, there were indications of divergence between clubs. However, this was not the case when using total property prices. By using simple statistical and graphical estimations, some important determinants for price growth and club formation seems to be geographic proximity, income, unemployment rate, debt gearing and market expectations.

The convergence formation in Oslo seems to be somewhat determined by geographic proximity. This can indicate that the geographic inequality and segregation in Oslo is further perpetuated by the development in the housing market.

Sammendrag

Målet med denne masteroppgaven var å analysere boligprismønsteret i Oslo, og å finne bevis for segmentering av boligmarkedet mellom de 15 bydelene. For å gjennomføre konvergenstestene brukte jeg den vektede repeterte salgsmodellen til Case and Shiller (1987) for å skape indeksen. Kvartalsvis data fra 1998 til 2019 ble brukt i denne analysen. En paneldatamodell skapt av Phillips og Sul ble brukt til å kjøre konvergenstestene. Dette ble gjennomført for både totale boligpriser og boligpriser per kvadratmeter.

Det ble ikke funnet noe bevis for at alle bydelene konvergerte. Med testing for totale boligpriser ble det identifisert tre konvergensklubber og en gruppe som divergerte. For boligpriser per kvadratmeter ble det også identifisert tre konvergensklubber, som dekket alle bydeler unntatt Alna.

Sett bort ifra ett unntak har alle gruppene holdt på sin posisjon i forhold til de andre gruppene. For gruppene som ble dannet med boligpriser per kvadratmeter var det indikasjoner for at konvergensgruppene divergerte fra hverandre. Dette var ikke tilfellet for klubbene dannet av totale boligpriser. Ved bruk av statistiske og grafiske estimeringer ble det identifisert noen forklarende variabler for prisutviklingen og klubbdannelse. Disse variablene ser ut til å være nærhet, inntekt, arbeidsledighetsrate, raten mellom gjeld og årlig inntekt og markedsforventninger.

Geografisk nærhet ser ut til å være en forklarende faktor for klubbdannelsene i Oslo. Dette kan indikere at økende geografiske ulikheter og segregering blir forsterket av boligmønsteret og utviklingen i dette markedet.

Content

List of Figures

List of Tables

1. Introduction

The housing market in Oslo has been subject of interest to the general population for some time now. The rapid price growth in the housing market in Oslo has been of particular interest in recent years. Since 2005, the housing prices per square meter in Oslo have increased by 171,31% in nominal values (Krogsveen, 2020, and based on data from Eiendomsverdi ASA). As a matter of fact, Oslo is the city with the fastest growing housing prices in Norway (Øye, 2019). Some of the explanatory variables for this price appreciation are high population growth, combined with an insufficient supply of properties. When demand increases faster than the supply, basic economic theory states that this increases the prices. It has been estimated that housing in Oslo was overpriced by 35% compared to the equilibrium price by 2012 (Krakstad & Oust, 2015, p. 19). People who have already entered the housing market in Oslo will appreciate this trend in growing property[1] prices. On the other hand, for those who have yet to buy a property for themselves, this trend is unfortunate, since it is harder to enter the market. Since properties such as houses and apartments are excellent saving objects, it is no wonder that many Norwegians are interested in the development in the housing market in Oslo.

It is not just the nominal property prices that are of interest. Other concerns in the Norwegian capital are the high socio-economic inequalities. In 2017, Oslo municipality had the second highest GINI-coefficient in Norway, right behind Bærum (Tuv, 2019). GINI is an index measuring income distribution inequality. Increased income inequality is a trend that seems to be growing. In 1986, the 10% richest earned 19% of the total income in Oslo. 14 years later, in 2000, the percentage increased to 27%. Also, the share of the population under 50% of the median income increased from 3,4% in 1986 to 7,6% in 2000 (Kirkeberg, 2003). There is also evidence for increased segregation in Norway based on findings from 1993 to 2017. Oslo is the most segregated city in Norway, both measured in income distribution and ethnic composition. These economic and ethnic disparities are also growing fastest in Oslo (Hernæs et. al., 2020).

[1] In this thesis, I will switch between using the terms *housing prices* and *property prices*. They mean the same thing.

Another study shows significant differences in life expectancy in the different districts in Oslo. The districts on the west side of Oslo have notably higher life expectancies than the districts on the east side (Dybendal & Skiri, 2005, p. 22). There is also a discrepancy for general criminality and violence within the city districts. Differences in the ethnic composition is also quite prevalent, where the eastern districts have a much higher share of an immigrant population than in the western parts of Oslo (Øia, 2007, p. 21). These differences and inequalities can have undesired effects, such as higher crime rates (Kelly, 2000), due to income inequality and possibly segregated communities, where sub-cultures are formed as a consequence of cultural and ethnic composition.

Aftenposten recently wrote an article about the "enormous" differences in the housing prices within Oslo, where the slowest price appreciation from 2015-2019 can be found in south-eastern parts of Oslo (Hager-Thoresen, 2020). Due to the aforementioned negative consequences of economic inequality and segregation, diverging property prices may perpetuate segregation and inequality and all the negative side consequences of economic and social inequality. Hence, the purpose of this paper is to analyse whether the housing prices in the districts have converged or diverged.

Price convergence is based on the theory that shocks or changes in one segment of the market has ripple-effects to other segments. This could also be called a catch-up effect. The theory is that prices in different segments will converge over time, which is intuitively sound. If the housing prices in one district increases, a possible reaction can be an increased demand for relatively cheaper housings in other districts. This thesis will cover an analysis for identifying overall property price convergence in Oslo or formation of convergence clubs in the districts within Oslo.

1.1 Hypothesis

The first hypothesis of this thesis is that there will be no evidence for overall convergence. In fact, there may even be divergent housing prices in Oslo.[2]

[2] The hypothesis can be formulated as: $\mathcal{H}_0: \delta_i \neq \delta \; for \; all \; i \; or \; \alpha < 0$. See Section 3 for detailed information.

The second hypothesis is that there can be identified convergence clubs within the different districts in Oslo.[3] Club convergence is defined as groups of districts in which the prices converges towards a steady state.

The third hypothesis is that the housing submarkets with relatively high prices have maintained the same position in the market as it did at the initial period. These tests will be conducted by using data collected from Eiendomsverdi AS.

1.2 Structure of thesis

This thesis will consist of six parts. The first section is the introductory part followed by the literature review in the second section. The third section is a detailed summary of the methodology used to conduct the convergence tests and how the index used in these tests was created. Section four is a brief overview of the data used in the thesis and its legitimacy. Section five consists of empirical results and discussion of the findings. The sixth and last part is the concluding remarks for this thesis.

1.3 Oslo

This section will provide background information about Oslo and the history of its housing market.

1.3.1 Districts

From 1988 to 2003 the number of districts were divided into 25 parts, but from 2003 and onwards the number of districts were reduced to 15 (Oslo kommune, n. d.). The current districts are shown in the map of Oslo below.

[3] The hypothesis can be formulated as: $\mathcal{H}_0\colon \delta_{ki} = \delta_k \; and \; \alpha \geq 0, where \; k \; is \; the \; club.$

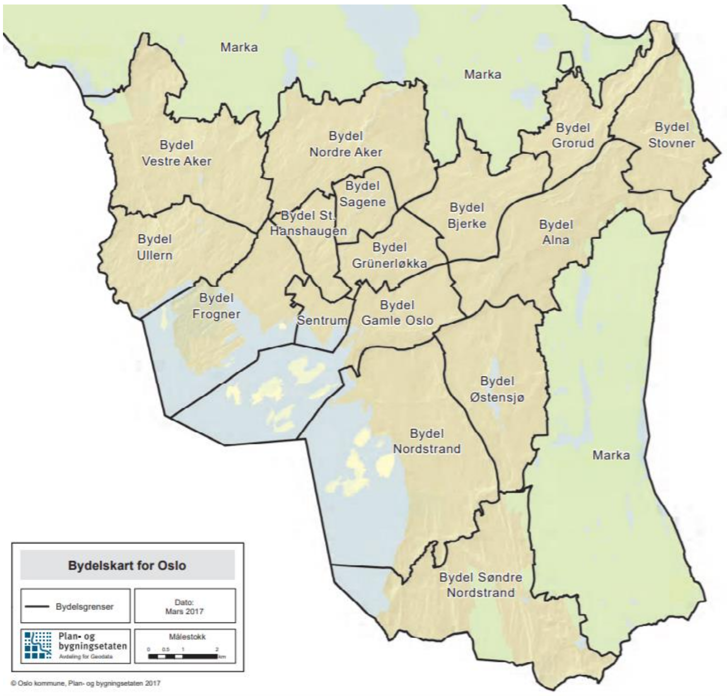

Figure 1.1: Map of Oslo and the city districts.

The 15 districts can be separated into five regions. In addition to the outer west, the inner west, the inner east and the outer east, the outer south is also a region suggested by Oslo municipality (Oslo kommune a, n. d.):

Outer West	Inner West	Inner East	Outer East	Outer South
Ullern	St. Hanshaugen	Gamle Oslo	Bjerke	Østensjø
Vestre Aker	Frogner	Grünerløkka	Grorud	*Nordstrand*
Nordre Aker		Sagene	Stovner	Søndre Nordstrand
			Alna	

Table 1.1: Overview of the 5 regions and their city districts in Oslo.

As mentioned earlier, there are quite a lot of economic differences in Oslo. This is also reflected in the housing prices. The western regions, as well as the inner east, are the most affluent areas with the most expensive housing in Oslo. Both the outer east and the southern region are, at large, the less affluent areas in the capital. This is also represented in the housing prices. One outlier is *Nordstrand*, which shares similar characteristics as the western regions, except its geographic placement.

1.3.2 History of housing market in Oslo

At the end of the 19_{th} century, the net migration to Oslo was at a massive level. Over the last 15 years, the population in Oslo increased with 70%. In 1898 alone, the population increased with 9%. This population growth was largely due a high economic activity. The housing prices in Oslo were also affected by this growth. From 1890 to 1899, the prices increased with 160% and the number of new residences was quadrupled (Lilleby, 2018).

Needless to say, Oslo became an attractive “hot spot” for housing speculants. However, this came at an abrupt end when the housing bubble Norway’s capital bursted, which is known as Kristianiakrakket. The property prices as well as the rent plummeted in 1899. In some cases, the cost of renting an apartment was zero as long as they kept living in the apartment. Approximately 12% of the appartments in Oslo were not occupied. It resulted in a steep decline in housing development. Furthermore, the unemployment increased rapidly. There were 28 000 registered as unemployed in Oslo in 1905 (Alsvik, 2009). This was despite of a massive spike in emigration, which quadrupled from 5 000 in 1898 to over 20 000 during Kristianiakrakket (Lilleby, 2018).

The population started to grow again in 1908. People were looking for working opportunities in the capital, but the willingness to build more houses and apartments was understandably quite low (Barstad, 2016, p. 6). In 1910, Oslo municipality decided to become a more active part of the housing development. Municipal housing development of social residences increased the overall building activity in the city. During the 20 years between 1911 and 1931, the municipality had built 7 200 apartments. Additionally, 6 300 apartments were built from the private sector with support from the municipality. From 1915 to 1929, the municipality became the largest owner of housing units in Oslo. However, after the 1920s, the municipality decided to stop the municipal housing construction and supported private construction instead. *Oslo og omegn Bolig- og Sparelag* (OOBS), today called *Oslo Bolig- og Sparelag* (OBOS) was founded in 1929 to build housing units to workers. From the period between 1936 to 1941, the number of newly constructed apartments with municipal support was 5 688 (Barstad, 2016, p. 10).

After the second world war ended, the demand for housing increased due to a wave of reimmigration. During the war, there was a decline in the population in Oslo, and the housing construction was at a halt. The municipality estimated that they needed 10 000 new apartments to combat the high number of homeless people, preferably with more than one- or two-room apartments. To help remedy this immediate housing crisis, they used the barracks from the Germans as housing units. Also, up until 1957, single people could not move into their own apartments (Barstad, 2016, p. 11).

In the 1960s, the housing market could be characterized as optimistic, with a high amount of constructions built as well as the formation of new satellite towns. From the 60s to the 70s, there was a large housing development in the satellite towns such as Manglerud, Oppsal, Tveita, Ammerud, Romsås and Holmlia. OBOS was credited for a lot of these constructions (OBOS, n. d.). In 1982, the price regulations on condominium were partly repealed, and housing market became close to regulated by market forces.

Deregulations in the financial market, combined with an economic boom, resulted in a high price appreciation in the housing market. The lending policy was an important driver for the economic growth. Even some of the consumption for the households were partly financed by loans (Torsvik, 1999).

In 1987, the bubble bursted, and the housing prices plummeted. When measured in real prices, the real estate prices between 1987 and 1992 fell with 43% in Norway. It was one of the largest housing price crashes in the history of Norway (Grytten, 2009).

The market began to stabilize in 1992. It had a stable growth path until the financial crisis in 2008. During the financial crisis, the housing prices fell with approximately 18% (Kutluay et. al., 2015). In Oslo, the prices have increased significantly since the economy stabilized after the financial crisis. Apart from some price drops in 2013 and 2017, the housing prices in Oslo have consistently increased (Krogsveen, 2020, and based on data from Eiendomsverdi ASA).

2. Literature

As far as testing for regional or national convergence, a large part of the research comes from the UK and USA, where they have found evidence of segmentation in UK (Montagnoli & Nagayasu, 2015; Abbot & Vita, 2013; MacDonald & Taylor, 1993; Cook, 2003) and USA (Kim & Rous, 2012; Kuketayev, 2013; Montañes & Olmos, 2013). They did not find any evidence for convergence between the regions overall, but convergence clubs were identified. Recently, similar results have been found in Australia (Awaworyi Churchill, 2018), China (Liri et.al., 2015), Poland (Żelazowski, 2019), South-Africa (Apergis et.al., 2015) and Turkey (Ganioğlu & Seven, 2019).

However, there have not been conducted that many convergence studies within a large city. A few examples are from Beijing (Gabrieli et.al., 2019) and Melbourne (Wong & De Silva, 2015). In Beijing, they found that over half of the property price differentials were converging. During economic growths, there was evidence of divergence between low- and high-price tiered properties in Melbourne. In other words, periods with high economic growth lead to a divergence in housing prices in the city.

2.1 Housing price indices

In order to conduct research on housing prices, the first step is to determine which price index is most suitable in the analysis. Some of the methods to estimate the price appreciation in the housing market are using the mean prices, the prices from repeated sales, hedonic sales and SPAR. This section will only cover the first three price estimation methods.

2.1.1 Average price index

This index is based on the average prices (alternatively the median price) on observed property sales, without controlling for heterogeneity. Furthermore, it does not take changes in the sample over time into account. The benefit with this approach is that it ensures many observations compared to other methods. Due to the sheer number of observations from this approach, it may be possible that the sample is somewhat comparable over time.

A typical weakness with this method is that new constructions may be larger, have more desirable characteristics, may be located at more attractive areas or have better services nearby. This can increase the average property size in the region, even if the prices for the rest of the housings remains constant. In other words, this approach makes little to no effort of adjusting for changes in the housing characteristics.

Aside from the sheer numbers of observations, there are no efforts to make sure the sample is representable of the actual housing stock. However, just as the simplicity of this approach definitely can lead to some very undesirable effects, it can also be considered a part of its strength, since it ensures a high number of observations. Still, the advantage of many observations is most likely heavily outweighed by the negative consequences of not adjusting for changes in the housing stock.

2.1.2 Repeat sales index

The second method is the repeat sales approach, where only repeated sales are used in the estimation. If the characteristics or quality of the properties remains the same, this method controls for heterogeneity. One potential weakness with this approach is that the observations of repeated sales are limited. The properties with repeated sales within a certain time interval may have qualities that differ from properties who only sold once, which can skew the estimation. This claim is supported by Case and Quigley (1991). They found that properties that were sold at least twice were much more expensive than the properties that were sold once. Case, Pollakowski and Wachter also found evidence that suggests housings that are sold more frequently appreciates faster (Case et. al., 1997).

Another potential problem is the constant-quality condition. To maintain a constant quality is not realistic, since most constructions will change over time. If properties are not maintained sufficiently, the quality will deteriorate. In this case, the appreciation of housing prices may be underestimated if constant quality is assumed (Harding et.al., 2007). An opposite scenario can also occur, since many owners spend a lot of time and money to improve their home environment. If this is just at the level where the quality remains constant, the repeat-sales index will be accurate. However, the appreciation will be overestimated if these upgrades increases the quality of the property, should a constant state be assumed. Abraham and Schauman (1991) and Peek and Wilcox (1991) showed how estimates for the 1970s and 1980s had an upward bias for the repeat sale index by 0.5% to 1% per year due to home improvement. Additionally, as Figure A2.1 in the appendix showcases, the expenditure per housing unit has increased quite a lot since the 80s. This can potentially increase the upward bias even more. Expenditure on maintenance and repairs have decreased since the 90s, but the opposite is the case for improvements. The latter is a probable driving force for higher property prices.

[*Figure A2.1 found in the appendix*]
[*Figure A2.2 found in the appendix*]

As we can see in Figure A2.2 (Prognosesenteret, 2014), Norwegians also seem to spend more and more money on renovation, rebuilding and extension. This is unfortunate for repeated sales-index, because this trend indicates a risk for overestimation of price appreciation as estimated by repeated sales in Norway.

2.1.3 Hedonic index

The third way is the hedonic method, which uses statistical techniques to control for heterogeneity. It attempts to estimate the value or price of certain attributes, which can be aggregated into the total price for a representative bundle of attributes. This is a solid way to control for heterogeneity.

A statistical regression will be run to identify the prices associated with the attributes to aggregate the estimated total housing price. These prices are based on how the observed housing prices are correlated with the attributes. When this is done, a representative bundle of attributes, which is based on the average quantity and/or quality of the attributes within a certain time period and region, is used to estimate the price index. However, this estimation method is prone to subjectivity. In other words, it can be difficult to determine the attributes just for a single property. Typical attribute options may be number of bedrooms, number of bathrooms, location services and centrality of the housing in question.

2.1.4 Comparison

Some researchers have attempted to compare how each methodology measures the appreciation of existing properties with constant quality. One study found evidence for repeat sales being a poor methodology when the time period is short (Clapp et.al., 1991). This was most likely due to the infrequent observations of repeated sales. Nevertheless, they found that if the periods extended to more than three years, the methodology became more and more accurate as the time period increased.

Another study was done by Crone and Voith (1992). When controlling for sampling size, they found that the repeat sales method holds up very well and had similar prediction errors as the hedonic method. In their comparisons, the average or the median approach was the clear loser, since its prediction errors was much worse than the two others.

Meese and Wallace (1997) had some important findings when they studied the price growth in Oakland Fremont in California in the 70s and 80s. Their panel data had over 20 000 observations, but when applying the repeat sales method, only 3 000 observations were available. Only 15% of the total observations were available for the repeated sales. They concluded that the appreciation estimated by the repeat sales method was likely too steep. Meese and Wallace (1997) stated that this was probably due to how repeat sales may bias the sample, especially with so few observations. The median and hedonic methods showed similar rates of housing price growth.

According to Rappaport (2007), there is no clear winner between the three methodologies. He compared three different indices to study the measured appreciation over time. The National Association of Realtors (NAR) is the average or median index, OFHEO HPI is the repeat sales index and the Census Constant Quality Index (CCQI) is the hedonic index. This is shown in Figure A2.3.

[*Figure A2.3 found in the appendix*]

Rappaport (2007) took note of some characteristics: i) Faster long-term growth for NAR, which was probably an upwards bias caused by the increased quality of the housing units. ii) Slowest growth for CCQI, which was a downwards bias caused by not controlling for the changed locations for the new homes. iii) The faster appreciation for the OFHEO HPI from 1999. Rappaport (2007) speculated that this was due to not controlling for the booming home renovation in this period. Moreover, properties where the price appreciates at a faster rate are more likely to be sold, which increases the likelihood of being included in the HPI index. iv) The HPI is the only index that appreciates smoothly over the time period. As a final conclusion, Rappaport (2007) found no clear winner for all scenarios and noted that all the methods have their strengths and weaknesses.

2.1.5 Chosen index

As shown in this section, it seems like no index is objectively superior under all circumstances. It depends on the panel data and on the nature of the analysis. The dataset for this thesis had initially over 300 000 observations, while the repeated sales had 122 063 observations. Consequently, the issue of few observations, as typically associated with the repeated sales method, is not of much concern for this thesis. Additionally, the time interval in this thesis is [1998Q1 - 2019Q4]. That is a much larger time period than the 3 years mentioned as the threshold for increased accuracy (Clapp et.al., 1991). Some of the weaknesses typically associated with the repeated sales method seems, for the most part, to be eliminated.

A hedonic index could also be suitable for running the convergence tests. However, since the repeated sales measures the price growth for the same housing units over time, it should be more robust for changes in the composition of the housing stock over time.

2.2 Convergence tests

There are a few commonly used methods to identify convergence in a data set. These methods will be briefly discussed below.

Granger causality test: One method of convergence testing is called the Granger causality test. This is combined with vector autoregression (VAR) models, impulse response and cointegration test (Cooper et.al., 2013). This method is mostly used for datasets with limited numbers of regions or countries. Due to constraints with the degrees of freedom, these models cannot use too many regions. According to Liri et. al. (2015), they usually contain about eight or less regional units.

Spatial/Temporal model: The spatial or temporal model is also used for convergence testing. This is a weighted model where variables from a neighbouring region are given more importance than variables from non-neighbouring regions (Gupta & Miller, 2012). The challenges with this approach are that the selection of spatial weights is a matter of subjectivity. This can create uncertainty in the estimation of convergence.

Unit root time series: A very common method for testing for convergence is the unit root time series. It was first implemented by Carlino and Mills (1993) in their seminal paper testing for convergence in regional per-capita income in the US. Even so, this method is not perfect, and has been criticized by Phillips and Sul (2007, 2009). If there is heterogeneity across regions or countries, this approach can be unsuitable. It cannot deal with this heterogeneity among individuals in the dataset. Another weakness with this method is that these tests can have a greater risk of over-rejection of the null hypothesis (Ng & Perron, 2001).

2.2.1 Phillips and Sul convergence test

In 2007 Phillips and Sul developed a panel data model to capture the behaviour of an economy in transition. It uses a log t test based on a linear regression to test for convergence. It consists of a *trend* component and *transitory* components. Before running the log t test, the transitory, alternatively called cyclical, component is filtered out. This econometric model addresses the problems with regular unit root time series, as mentioned above. It takes heterogeneity of individuals into account.

Another benefit with their model is that it enables testing for regional clusters. In other words, it makes it possible to test for convergence clubs. This is particularly useful, since rejecting the null hypothesis of overall convergence does not mean that there is no convergence in the regions at all. There may still be districts where the housing prices clusters towards a steady state (Phillips & Sul, 2007). Due to these benefits, this is the model that will be used in this thesis.

2.3 Ripple effect

The ripple effect refers to the phenomena of a “domino”-effect in the housing market. Typically, the case from Britain (Meen, 1999) is often used as an example of this phenomena. The prices initially rise in London or South-East Britain, and this effect spread out to the rest of the British island.

There are different theories that attempts to explain this effect. One theory is that this pattern is reflected by the economic growth rates within the regions. Another popular explanation is migration patterns. This directly translates into changes in demand in the market. If one region has become expensive, a natural reaction to this is increased demand in relatively cheap areas.

Still, as shown by Meen (1999), this theory has a weak explanatory power, at least on its own. Meen argued that higher debt ratios in the southern regions makes said region more sensitive to changes in economic indicators such as interest rates, wealth and unemployment. The ripple effect may be a good reason as to why the property prices in some regions may converge. If this theory holds true, the ripple effect can somewhat explain how the prices in different housing markets converge towards a steady state.

2.4 Drivers for price appreciation

In simple terms, the housing market, just as any other market, must adhere to some fundamental economic principles such as demand and supply. This section will cover some of the price determinants for real estate.

2.4.1 Aggregate demand for housings

According to Natsvaladze and Beraia (2018), the exogenous variables for real estate demand is market size, which includes factors such as *population* and *employment* (more relevant for office spaces), *economic status*, which means income or wealth, *the alternative prices* (substitutes) in another market segment, *expectations* of market shocks or changes, *interest rate* and access to *credit*. Natsvaladze and Beraia (2018) suggested that a useful way to measure the demand in practical terms is to use the term net absorption. This is defined as changes in a market's *occupied stock* between a particular time interval.

2.4.2 Aggregate supply in the housing market

Natsvaladze and Beraia (2018) suggests the short-term supply curve is completely, or at least close to, inelastic. Therefore, it is not feasible to increase the supply of housing units with an immediate notice. The construction duration is the cause of this short-term inelasticity. The construction lag in the US is, according to Natsvaladze and Beraia (2018), somewhere between 6-12 months for housing units. Short-term in the housing market, as suggested by Kongsrud (2000), should be seen as a 2 - 3 year period. So, within this time period, the supply curve will be completely inelastic, and has therefore little sensitivity to changes in demand.

This is not the case from a long-term perspective, as the supply curve becomes more elastic over time. Another bottleneck for production of real estate is the general space required for construction, capital, labor and building materials. Policy-limitations, such as receiving permit to construct new real estate can also be considered a bottleneck. Natsvaladze and Beraia (2018) also suggested that *the production costs*, *the subjective market risk*, *expectations of coming real estate prices* and *availability* affected the supply in the housing market.

2.4.3 Determinants in Norway

Statistisk sentralbyrå (SSB) has developed a macroeconomic model (MODAG) for the Norwegian economy. MODAG is mainly used by the Norwegian Ministry of Finance, but SSB is also using it for their own analysis. According to this model, the demand for buying a property in Norway is dependent on *the housing price*, *the disposable real income for the household* and *real interest after taxes*. The housing prices are related with *income*, *real interest*, *housing stock* and *new construction*. Supply is determined by the explanatory factor *stock of existing housing stock*, which is changing over time due to investment. Investments are affected by *the price of existing properties* and *building costs* (Baug & Dyvi, 2008).

3. Method

3.1 Overall convergence

To determine if the property prices in Oslo have converged over the last 20 years, the econometric model created by Phillips and Sul (2007, 2009) will be applied in this thesis. It is a panel data model, which means the data will be divided into individuals (i) and time (t). Panel data is usually constructed as this:

(1) $$X_{it} = c_{it} + a_{it}$$

In the function above, X_{it} is the data variable. In this thesis, this represents the log average property price of region i in period t, or simply the panel data variable. c_{it} represents the systematic components, which also includes permanent common components. This can also be called the trend component. a_{it} covers the transitory components. At this point, equation (Eq.) 1 can consist of both common and idiosyncratic components in both the parameters. Phillips and Sul (2007) suggested to transform Eq. 1 in order to separate the idiosyncratic and common components in the following way, for $i = 1, \ldots, N$ and $t = 1, \ldots, T$:

(2) $$X_{it} = \left(\frac{c_{it} + a_{it}}{\mu_t}\right)\mu_t = \delta_{it}\mu_t, \text{ for all } i \text{ and } t.$$ [4]

In the equation above, μ_t represents the single common component. δ_{it} is the idiosyncratic parameter that varies over time. To explain it in simpler terms, δ_{it} is a measure of the percentage in μ_t of individual i at period t. In other words, it measures the difference between μ_t and X_{it}

The next step is to make it possible to estimate δ_{it}. By making some restrictions on δ_{it} and μ_t, it is possible to remove the common factor μ_t with the following regression:

(3) $$h_{it} = \frac{X_{it}}{\frac{1}{N}\sum_{i=1}^{N} X_{it}} = \frac{\delta_{it}}{\frac{1}{N}\sum_{i=1}^{N} \delta_{it}}, X_{it}, \delta_{it} > 0.$$ [5]

[4] Phillips & Sul, 2007, p. 1774.
[5] Phillips & Sul, 2007, p. 1780.

This function measures δ_{it} relative to the average in the panel at period t. h_{it} is thus called the *relative transition parameter*. It identifies a transition path of region i relative to the panel average at period t. An assumption made by Phillips and Sul (2007) is that $N^{-1}\sum_{i=1}^{N}\delta_{it}$, the panel average, will almost definitely differ from zero as N → ∞. Also, X_{it}, δ_{it} and μ_t in my dataset are all positive, which means that the structure of Eq. 3 will not cause any practical issues. Eq. 3 has two noteworthy properties. The first property is that the cross sectional mean of h_{it} is unity. The second property is if

(4) $\lim_{t\to\infty}\delta_{it} = \delta$, for all i and j

, h_{it} converge to unity. As $t \to \infty$, or in more practical terms, in the long run, the variance of h_{it} converges towards zero. A requirement for convergence is:

(5) $\lim_{k\to\infty}\frac{X_{it+k}}{X_{jt+k}} = 1$, for all i and j.[6]

As Phillips and Sul (2007) pointed out, this can be defined as the relative convergence. An observant reader may also notice that the convergence in Eq. 4 and Eq. 5 are equivalent to each other. The semiparametric form δ_{it} that allows transitional heterogeneity.[7] This means that even if $\delta_i = \delta_j$, transitional periods, or $\delta_{it} \neq \delta_{jt}$, may still occur. An assumption we make is that the idiosyncratic parameter, δ_{it}, is:

(6) $\delta_{it} = \delta_i + \sigma_{it}\xi_{it},\ \sigma_{it} = \frac{\sigma_i}{L(t)t^{\alpha}}, t\ \geq 1, \sigma_i > 0$, for all i[8]

Coefficient ξ_{it} is weakly dependent over t. Also, it is required that it is *iid*(0,1) for every *i*. The function in the denominator, *L(t)*, is a function that varies slowly. This function can be constructed in different ways – *log(t)*, $log^2(t)$ or log(log(t)). Phillips and Sul (2007) has recommended to set $L(t) = \log(t)$, since this function has smaller size distortions, and it also has the best testing power[9]. *L(t)* is increasing as *t* increases, and it is divergent as it goes towards infinity. This means that convergence as defined by Eq. 4 and Eq. 5 is dependent on

[6] Phillips & Sul, 2007, p. 1779.
[7] Phillips & Sul, 2007, p. 1773.
[8] Phillips & Sul, 2007, p. 1785.
[9] Phillips & Sul, 2007, p. 1803.

the value of a, the decay rate. As shown by Phillips and Sul (2007), convergence is occurring when $a \geq 0$. The null hypothesis can thus be formulated like this:

(7) $\mathcal{H}_0: \delta_i = \delta \; and \; \alpha \geq 0$[10]

The alternative hypothesis of no convergence is shown below:

(8) $\mathcal{H}_A: \delta_i \neq \delta \; for \; all \; i \; or \; \alpha < 0$

To test this hypothesis of convergence, Phillips and Sul (2007) created a *t* test which can be applied by the following *log t* regression:

(9) $log\left(\frac{H_1}{H_t}\right) - 2\log(\log(t)) = a + blog(t) + \varepsilon_t$

$for \; t = [rT], [rT] + 1, \ldots, T$, where $r > 0$[11],

In the regression above, $H_t = \frac{1}{N}\sum_{i=1}^{N}(h_{it} - 1)^2$. Phillips and Sul (2007) also proved that $b = 2a$. The null hypothesis is a one-sided test of $a \; or \; b \geq 0$. Long-run convergence is occurring if $log\left(\frac{H_1}{H_t}\right) \rightarrow \infty$ as $2\alpha\log(t)$ when $a > 0$.

We reject the null hypothesis when $t_b < -1.65$ at a 5% significance level. The asymptotic theorem for the t_b distribution is:

(10) $t_b = \frac{\hat{b}-b}{s_b} \Rightarrow N(0,1)$[12]

The value selected for r can affect the results from the null hypothesis test in Eq. 9. A satisfactory performance is ensured when $r \in [0.2, 0.3]$, as shown by the Monte Carlo experiments. The selection is dependent on the amount of time periods there are in the panel.

---

[10] Phillips and Sul, 2007, p. 1788. For convergence clubs, the hypothesis can look like this: $\mathcal{H}_0: \delta_{ki} = \delta_k \; and \; \alpha \geq 0, where \; k \; is \; the \; convergence \; club.$

[11] Phillips & Sul, 2007, p. 1789.

[12] Phillips & Sul, 2007, p. 1790.

Phillips and Sul (2007) suggested that when it is a small or moderate T *($\leq$ 50)*, we should set $r = 0.3$. If it is a large T*($\geq$ 100)*, the selection should be $r = 0.2$.

3.2 Club convergence

If we reject the null hypothesis of overall panel convergence, this does not necessarily mean there is no convergence in the panel at all. One region can be split into several subgroups and tested to identify any cases of equilibria or steady state growth paths. There may also exist clusters that diverges in the panel data. Philips and Sul (2007) created a way to test for club convergence. This section will briefly provide a short summary of the five steps used to identify convergence clubs.

Step 1: Last Observation Ordering. The first thing we do is to sort the individuals, or in this case, the regions, in the panel with accordance to the last observation. If the time series is substantially volatile in X_{it}, we can sort the regions based on the time series average,

(11) $(T - [Ta])^{-1} \sum_{t=[Ta]+1}^{T} X_{it}$ [13],

over the last fraction $(f = 1 - a)$ of the panel data. Examples for this can be $f = \frac{1}{3}$ or $\frac{1}{2}$.

Step 2: Core Group Formation. After *Step 1*, we select the *k* highest regions in order to create the core subgroup G_k, with the following conditions:

(12) $N > k \geq 2$

Then we run the *log t* regression to test for convergence, $t_k = t(G_k)$, in the subgroup. The goal is to find the k^* that has the highest t_k-level. Consequently, the conditions are:

(13) $k^* = arg\max_{k}\{t_k\} \quad subject\ to \quad \min\{t_k\} > -1.65.$

The condition above is important to make sure the null hypothesis is valid for each k. One thing to be aware of, as with all null hypothesis testing, is the threat of not rejecting a false

[13] Phillips & Sul, 2007, p. 1800.

null hypothesis, which is also called a $type\ II$ error. Finding k^*, as shown in Eq. 13, is a way to reduce the probability for $type\ II$ error to occur. Furthermore, if there are no core convergence subgroups that satisfies our conditions from Eq. 13, there is no evidence of any convergence clubs in the panel. If we do find one that satisfies the condition, this group can be denoted as G_{k*}.

Step 3: Sieve Districts for Club Membership.[14] The next step is to create a complementary set, G_{k*}^c, to the core group created in *Step 2*, G_{k*}. This complementary set consists of individuals who are not in G_{k*}. We extract one individual from G_{k*}^c to the core convergence subgroup. Then we run the *log t* test to see if we should include said individual in the core subgroup. The *t*-value from this test, $\hat{t}$, determines if we should include the individual or not into the initial group. If $\hat{t} > c$, where c is defined as the chosen critical value, it should be added into the core subgroup to form a new convergence group. Since *T* has many observations, the critical value, c^*, Phillips and Sul (2007) recommended, through Monte Carlo testing, to set the critical at -1.65 when *T* is not small.

Step 4: Stopping Rule.[15] Now we form a new group of the individuals who were not sieved in *Step 3*. In other words, the individuals who had $\hat{t} < c$ will now be a part of the new subgroup. Just as done before this, another *log t* test will be run to see if $t_b > -1.65$, which means the new cluster converges. If this is the case, the panel consists of at least two convergence clubs. If this criteria for convergence is not upheld, we will repeat *Step 1-3* again to see if there are any remaining subgroups with some convergence. Naturally, we repeat this procedure until $t_b > -1.65$ for the individuals who have not yet been sieved, or until we cannot find any k where $t_k < -1.65$. If the latter holds true, the conclusion is that the rest of the individuals or regions diverges.

Step 5: Club Merging. In this step, we run *log t* tests for all combinations of pairs of the initial convergence clubs identified in all the prior steps. If any of these pairs shows evidence for convergence, these pairings will be merged to form a new convergence club. Schnurbus et. al. (2016) made some suggestions to run club merging tests. First, a *log t* test will be run for Club 1 and Club 2. If these newly merged clubs converge, this will be the new Club 1.

[14] Phillips & Sul, 2007, p. 1801.
[15] Phillips & Sul, 2007, p. 1801.

The next step is to merge the newly formed Club 1 with Club 3, and then run the *log t* test for this club combination. If the null hypothesis of convergence is rejected for initial Club 1 and Club 2, a *log t* test for convergence will be ran on the pairing of initial Club 2 and Club 3. This process will be repeated with all the initial clubs. Also, this process can also be applied to the newly founded clubs until there are no more clubs to merge.

3.3 Stata commands

This section will provide a brief explanation and summary of the Stata commands that will be used to test for convergence in this thesis. Credit for these commands goes to Kerui Du (2018), who developed a new Stata module with five commands to run the aforementioned convergence tests made by Phillips and Sul (2007). For more detailed syntax explanations, Du (2018) has provided a thorough explanation if necessary.

Before using any of the commands provided by Du (2018), some data preparation must be made. The data provided by Eiendomsverdi included all the transactions within the given time period and for all periods. These transactions must first be aggregated to the mean property price for i in t. This is necessary, because in order to be able to filter components in a time series, each X_{it} must be unique. In this panel, the time series will be *quarterly* split up for each year ($YYYYQ$). Another small adjustment is to transform the property prices into log values, since this method requires a log transformed X_{it}. It is also necessary to declare the data as a panel, by using the $xtset$-command, followed up by i- and t-variables.

3.3.1 Filtering components

As explained in section 3.1, one of the first things that needs to be done in order to run the *log t* regression for the convergence test is to filter out the cyclical component. The $pfilter$ command developed by Du (2018) is used to filter out the cyclical component. The command has 4 options as to which filtering system that will be used. This thesis will make use of the filtering method developed by Hodrick and Prescott (HP) (1997). Filtering components by using the HP-method is a popular choice for many researchers, due to its flexibility and simplicity (Phillips & Sul, 2007, p. 1783).

This is selected by typing $method(hp)$ in Stata as an option. As mentioned, the goal is to wipe out the *cyclical component*. This can be done by typing $trend(newvarname)$, which is an option to store the *trend component*. The last option is the smoothing parameter, which smoothens out the nonstationary trends in the time series. The default rule of thumb is to use a smoothing parameter equal to $1\ 600$ when dealing with quarterly panels (Hodrick & Prescott, 1997), and this is the smoothing parameter that will be used in this thesis. In Stata, this is done by plotting $smooth(1600)$ as an option.

3.3.2 Log t test

To run the *log t* test in Stata, the next step is simply to plot in $logtreg\ (varname)$, where $varname$ is the trend component that were filtered out in section 3.3.1. This regression has heteroskedasticity- and autocorrelation-consistent standard errors. It also gives us the option to what proportion of the data that should be discarded before running the regression, by plotting $kq(\#)$ in Stata. This command is used in the remainders of the commands. The following result will either be a rejection or non-rejection of the null hypothesis of overall convergence.

3.3.3 Club convergence

To run the *log t* test for club convergence, we plot in $psecta\ varname$, where $varname$ is still the trend component filtered out in the filtering process. There are many available options provided by Du (2018). One important option is to set the critical value for club clustering. Phillips and Sul (2007) showed that when $\alpha = 0.2$, which is the critical value, the rate of $Type\ I$ & $Type\ II$ errors will be smaller than when $\alpha = 0$. Therefore, the critical value in this thesis will be 0.2. This is done simply by plotting $cr(0.2)$ as an option in Stata. It is also necessary to store the new $club$ variable in the panel, which is required to do further testing for club merging. Fortunately, there is an option command provided by Du (2018), where you simply type in gen$(varname)$ in Stata. These are the initial club classifications.

3.3.4 Club merging

It is quite easy to see if there are any clubs paired together that converges. The command in stata is $scheckmerge\ varname$, where $varname$ is the trend component. This is followed by the option where you specify the initial club subgroups, $club(varname)$, that is used for pairing adjacent initial clubs. If there are any pairings that converges, we can use the command $imergeclub\ varname$. Yet again, $varname$ is the trend component filtered out from the beginning. The same initial clubs are used in this command, so $club(varname)$ is just the same as it was for $scheckmerge$. The final part of this process is to generate the newly formed clubs, which is done by plotting $gen(newvarname)$. After running this command, the final results of the convergence testing will be provided.

3.4 Case & Shiller weighted repeat sales method

As mentioned in section 2.1, it is important to control for heterogeneity when creating a housing price index. This section will therefore outline how to create a repeated sales index to control for changes in the housing market.

The weighted repeat sales (WRS) is a modified version of the standard model created by Bailey, Muth and Nourse (BMN) (1963). Their model is created by subtracting the log-value of the first sale from the log-value of the second sale for property i. The dependent variables are only made up by dummy variables, which are used to identify the period the property was sold. The structuring of the model can be formulated as shown by Eq. 14.

$$\log(P_{it}) - \log(P_{is}) = \delta_2 D_{i2} + \delta_3 D_{i3} + \cdots \delta_m D_{im} + \varepsilon_{it}, \quad (14)$$
$$i \in I; t, s \in \{2, \dots, m\}, D_{it} \in \{-1,0,1\}$$
$$s < t$$

The dependent variable, p, is the sale price. The dummy, D, will identify first sales, no sale and second sale. First sale is denoted by the time period s, while the second sale is denoted by t. As shown in the model above, this means that s is smaller than (before) t. A repeated sale is denoted by i, and I is the subscript for all the repeated sales in the panel. Parameter δ is the index to be estimated. The error term, ε_{it}, has a zero mean and has constant variance. The dummy variable is either denoted as $\{-1, 0, 1\}$. The first sale sets $D = -1$, the second sale is $D = 1$, while no sale means that $D = 0$. If the first sale is in the first period in the time interval [0,T], the dummy will be set as 0, not -1. This is because the index in the first period will be 1. Finally, the time interval [0,T] is split up into m parts.

If the error term is independently normally distributed with zero mean, the BMN-model is sufficient to ensure minimized variance and unbiased estimators of the δ-coefficient. However, as argued by Case and Shiller (1989), the error term is often related to the time interval between s and t, and the errors increases as the interval increases. In other words, the errors are heteroskedastic. This is also the case in this thesis, as shown in Table A3.1 in the appendix.

[*Table A3.1 found in the appendix*]

Table A3.1 shows that the time interval between sales is positively correlated with the residuals in Eq. 14, which indicates heteroskedasticity. To remedy this, they created a WRS-index, where less weight is given to long time intervals. Their model consists of three steps:

i) Simply run the regression in Eq. 14.

ii) After running the regression in step 1, we must predict the residuals from said regression. Then run a regression of the squared residual as the dependent variable, and with a constant term and the time interval $[s, t]$ as independent variables. The function in the second step is as following:

(15) $\varepsilon_{it}^2 = c + TI_{it} = \beta_0 + \sum_{t=1}^{I}(\beta_1 HP_{it}) + \mu_{it}$

Where c is the constant term, μ_{it} is the error term for Eq. 15 and TI is the time interval $[s, t]$.

iii) After running the regression in step 2, we run a weighted generalized least squares regression as done in Eq. 14. However, this time, all the observations are divided by the square root of the fitted value from Eq. 15.

The following result will show the coefficients for each time period. To complete the WRS-index, we return e to the power of coefficient δ from the weighted Eq. 14.

(16) $I_{it} = e^{\delta_{it}}, t \geq 1$

In this case, i is a parameter for the city districts and I is a denotation for the WRS-index. After running Eq. 16, the WRS-index is completed, and it shows the price appreciation since the base year. In this thesis's case, the base period is year 1998 in the first quarter.

3.4.1 Setting up the observation pairs

In the panel there are many cases where properties have been sold more than twice. In order to ensure a comfortable margin of observations, these property sales will also be included. By including properties that are sold more than twice, this will increase the likelihood of selecting units with special characteristics, i.e. more expensive or faster appreciation rate. Nevertheless, the benefit of selecting for more observations probably outweigh the possible selection bias. To include them in the WRS-index, we simply have to treat each two sales as one observation. In other words, if there are more than two sales for housing i, the first two sales will be treated as the first pair, while the second and third sale will be treated as the second pair and so on until there are no more sales for said property.[16]

3.4.2 Implementing repeated sales in the convergence test

To use this in the convergence test provided by Phillips and Sul (2007), some further steps must be done. The index described in section 3.4 only shows the price appreciation, where the index for all cities starts at 1 at the base year. This means they have the same starting point, and this will skew the convergence test. Fortunately, this is easy to fix by simply using the panel for repeated sales to estimate the average price for the base year. The next step is just to multiply this average price for all the districts with their respective price appreciation as estimated by the WRS. These simple adjustments will make it possible to use WRS for the convergence tests created by Phillips and Sul (2007). The final transformation looks like the indexing below.

$$P_{it} = I_{it} \times P_{t=0,i} \tag{17}$$

[16] This approach makes it tough to argue that the observations are *independent and identically distributed* and will not be solved by using robust standard errors.

4. Data

4.1 Validity

This thesis uses a WRS-index following the Case & Shiller-method (1989). Due to limitations with the number of observations, the convergence tests will be run for all types of residences. It would be good to, for instance, run the tests for apartments only, but there are not enough observations for this. Only using apartments for in this analysis leads to “holes” in the index. The low number of observations for each period and district would lead to weak regressions.

Another issue with only using apartments is that the coefficients in the WRS-index suffered from low statistical significance. Running the convergence tests for all the housing types can be considered somewhat of a weakness in the analysis, since the housing types varies somewhat between the city districts. Other than that, the dataset and index should be quite solid, especially since there are 121 614 repeated sales between 1998 to 2019. It would be interesting to use to a hedonic price index in this thesis. However, creating such an index is quite demanding, both measured in time and available data, and is therefore beyond the scope of this thesis. This should not be an issue, since the WRS-index should sufficiently take heterogeneity into account.

4.2 Quality of data

The data used in this thesis is collected from Eiendomsverdi, who provides a property transaction database, as tasked by Eiendom Norge. Eiendomsverdi is a company that was founded in year 2000, and they have a database of property transactions going back to 1990 (SpareBank, n. d.). They have developed a hedonic SPAR-index based on data from 2003 (Eiendom Norge, n. d.). Hence, there are limitations with how far back you can go with a hedonic index. This is yet another reason for why the WRS-index is beneficial for this thesis. It allows me to create an index from further back than 2003, which I have used to analyse the data from 1998 to 2019.

When a residential sale has been made, the housing price will be updated the day after in their databank. This makes it possible to create both an average- and repeat sales index. The data and statistics provided by Eiendomsverdi is used by many important institutions. It is used by many researchers and for analytical purposes, news companies and likely by political groups and governing bodies, such as Norges Bank (Eiendomsverdi, n. d). Eiendomsverdi can therefore safely be considered as a trustworthy and respected data collector and provider. Credibility like this will increase the quality of the data, and it increases the reliability of this thesis.

4.3 Overview of data

As we can see in Table 4.1, which is a summary statistic for the housing prices in Oslo and the data set in general, the district with the highest mean prices in the period from 1998q1 to 2019q4 is Vestre Aker, followed by Nordstrand and Ullern. The three lowest mean prices in Oslo can be found in Bjerke, Søndre Nordstrand and Stovner respectively. A fortunate characteristic of this dataset is that there are thousands of observations in each suburb. However, the interval between the lowest and highest amount of observations $[Grorud, Frogner;\ 3034, 18363]$ differ by quite a large margin. In any case, the sample size should be sufficient for the purpose of a statistical analysis.

District	*Mean Ln Price*	*Min Ln Price*	*Max Ln Price*	*Std. Dev.*	*Observations*
Vestre Aker	15.386	14.493	16.076	0.436	6101
Nordstrand	15.283	14.318	16.02	0.472	7028
Ullern	15.275	14.421	15.962	0.437	5235
Nordre Aker	15.159	14.134	15.897	0.496	5602
Frogner	15.154	14.188	15.871	0.463	18363
Østensjø	14.942	13.982	15.689	0.478	7102
St.Hanshaugen	14.855	13.87	15.587	0.478	11182
Alna	14.703	13.855	15.421	0.439	6148
Grünerløkka	14.658	13.533	15.422	0.495	15368
Gamle Oslo	14.623	13.588	15.376	0.481	11036
Grorud	14.607	13.65	15.35	0.461	3034
Sagene	14.566	13.48	15.358	0.414	12309
Stovner	14.556	13.725	15.225	0.516	3525
Søndre Nordstrand	14.543	13.754	15.167	0.393	4463
Bjerke	14.514	13.571	15.256	0.473	5118

Table 4.1: Summary statistics. The natural logarithm of the prices is transformed from the total housing prices in NOK.

In Table 4.2, we see that the three most expensive districts are Frogner, St. Hanshaugen and Sagene respectively when measured in prices per square meter. At the opposite end, the three least expensive districts are Stovner, Søndre Nordstrand and Alna respectively. If we compare Table 4.1 with Table 4.2, it is worth mentioning that the districts position relative to others differs from the total housing price and housing price per square meter. This is why it is of interest to run convergence tests for both total prices and prices per square meter.

Bydel	Mean Ln Price	Min Ln Pris	Max Ln Price	Std. Dev.	Observations
Frogner	10.707	9.738	11.424	0.463	18363
St. Hanshaugen	10.617	9.628	11.35	0.478	11182
Sagene	10.573	9.487	11.366	0.516	12309
Grünerløkka	10.552	9.419	11.317	0.495	15368
Nordre Aker	10.529	9.503	11.268	0.496	5602
Ullern	10.514	9.658	11.202	0.437	5235
Vestre Aker	10.503	9.609	11.193	0.436	6101
Gamle Oslo	10.474	9.44	11.228	0.481	11036
Nordstrand	10.421	9.456	11.158	0.472	7028
Bjerke	10.359	9.404	11.101	0.473	5118
Østensjø	10.274	9.314	11.022	0.478	7102
Grorud	10.236	9.279	10.979	0.461	3034
Alna	10.166	9.317	10.884	0.439	6148
Søndre Nordstrand	10.028	9.238	10.653	0.393	4463
Stovner	9.961	9.117	10.63	0.414	3525

Table 4.2: Summary statistics. The natural logarithm of the prices is transformed from the housing prices per square meter in NOK.

5. Results

This section will consist of different types of convergence tests. Results for total housing price and the price per square meter will be provided.[17] These tests will be conducted by using a WRS-index. An average index will also be used for comparing what happens if heterogeneity is not controlled for.

5.1 Overall convergence for total housing prices with WRS

In Table 5.1, we reject the null hypothesis of convergence for overall convergence in Oslo. This is because $t < -1.65$ by quite a margin ($t = -17.7581$). This means that, by using the WRS-index, there is no statistical evidence for overall convergence in the entirety of Oslo.

Variable	Coefficient	Standard Error	T-Statistic
Ln Housing Price	-0.3829	0.0216	-17.7581*

*Table 5.1: Convergence test for overall convergence in Oslo measured in total housing price. The number of districts is 15. The number of time periods is 88. The first 26 periods are discarded before regression. * is denoted as a rejection of the null hypothesis of convergence at 5% significance level.*

This does not mean that there is no convergence in Oslo. To test if there are any evidence of convergence, the next step is to test whether housing prices converges to some steady states within soubgroups. In other words, are there any group of city districts in Oslo that forms convergence clubs?

log(t)	Club1	Club2	Club3	Club4	Group5
Ln Housing Price	0.0525	0.8366	0.2852	2.1567	-0.9234
T-stat	1.4859	36.6804	1.1155	1.0747	-92.3234*

*Table 5.2: Convergence test for club convergence in Oslo. * is denoted as a rejection of the null hypothesis of convergence at 5% significance level.*

[17] The price per square meter is measured with usable area (Bruksareal/BRA). This is the area of all rooms that are at least 1.9 meter tall and 0.6 meter wide at said height. It is measured from the inside of the walls (Norsk takst, 2015). It only measures the inside of the building, which means that balconies and terraces are not included in the measurement.

Convergence club	*Districts*
Club 1	| Frogner | Nordre Aker | Nordstrand | Ullern | Vestre Aker | | Østensjø |
Club 2	| Alna | Grünerløkka | Sagene |
Club 3	| Gamle Oslo | Grorud |
Club 4	| Bjerke | Stovner |
Group 5	| St. Hanshaugen | Søndre Nordstrand |

Table 5.3: Overview of the initial convergence clubs.

Table 5.2 and Table 5.3 shows that there, at this point, four convergence clubs and one divergence club in Oslo. All the clubs are converging, except for Group 5, which consists of St. Hanshaugen and Søndre Nordstrand. This is because the $t < -1.65$ by a huge margin (-92.323). The others all have t-values higher than -1.65, which is evidence of convergence. It is possible that the formed clubs can be merged with adjacent steady-state clustering clubs, which will be the next step in these tests.

log(t)	**Club1+2**	**Club2+3**	**Club3+4**	**Club4+G~5**
Ln Housing Price	-0.1978	0.2810	-0.9835	-0.7118
T-stat	-6.9122*	4.0805	-21.0787*	-58.9159*

*Table 5.4: Test for initial convergence club merging with adjacent clustering clubs. * is denoted as the rejection of the null hypothesis of club merging at 5% significance level.*

As Table 5.4 shows, there is only one pair of clubs that clusters with a 5% statistical significance level. In the rest of the tests, we reject the null hypothesis of club convergence. Since two of the initial clubs could be merged, the new list of convergence clubs is listed in Table 5.5.

log(t)	**Club1**	**Club2**	**Club3**	**Group4**
Ln Housing Price	0.0525	0.2810	2.1567	-0.9234
T-stat	1.4859	4.0805	1.0747	-92.3234*

*Table 5.5: The new convergence clubs after merging initial Club 2 and Club 3. * is denoted as the rejection of the null hypothesis of club merging at 5% significance level.*

Since two of the initial clubs converged, we should consider that further club merging can be possible. As shown in the Table 5.6, no clubs can be merged anymore, so the clubs formed in Table 5.5 remains. The final clubs are shown in Table 5.7.

log(t)	Club1+2	Club2+3	Club3+G~4
Ln Housing Price	-0.2430	-0.7482	-0.7118
T-stat	-9.5777*	-13.5452*	-58.9159*

*Table 5.6: Test for merging clubs a second time. * is denoted as the rejection of the null hypothesis of club convergence at 5% significance level.*

Convergence club	*Districts*
Club 1	| Frogner | Nordre Aker | Nordstrand | Ullern | Vestre Aker | | Østensjø |
Club 2	| Alna | Gamle Oslo | Grünerløkka | Grorud | Sagene |
Club 3	| Bjerke | Stovner |
Group 4	| St. Hanshaugen | Søndre Nordstrand |

Table 5.7: Overview of the final convergence clubs.

Club 1 mainly consists of districts in inner and outer west. However, Nordstrand and Østensjø are also in the club. The former district is not surprising, since Nordstrand have more in common with the western boroughs. Figure A5.1, Figure A5.2 and Figure A5.3 in the appendix shows how the unemployment rate and ethnic composition is quite similar in the districts in Club 1.

[*Figure A5.1 found in the appendix*]

[*Figure A5.2 found in the appendix*]

[*Figure A5.3 found in the appendix*]

The districts in Club 2 are all in the inner or outer east, and they are also mostly neighbouring areas. Club 3, or Bjerke and Stovner, are quite close to each other, and they are both in the outer east of Oslo. Also, their income level, unemployment rate and ethnic composition are all quite similar. Group 4, or St. Hanshaugen and Søndre Nordstrand, are in different parts of Oslo. They are also somewhat different when it comes to particularly unemployment rate and the ethnicities in the respective districts. However, their income level and debt gearing are not too different from each other. With Club 2 as an exception, Figure A5.4 shows that the debt gearing is somewhat similar within the convergence clubs.

[*Figure A5.4 found in the appendix*]

All in all, it seems like convergence or divergence is somewhat dependent on similar or dissimilar socioeconomic characteristics and proximity. The latter may be explained by some sort of local ripple effect in the neighbouring areas, as can be seen in the map of Oslo depicting the convergence in Figure 5.1. Such a local ripple effect can indicate that there are certain geographic amenities, such as in the clubs.

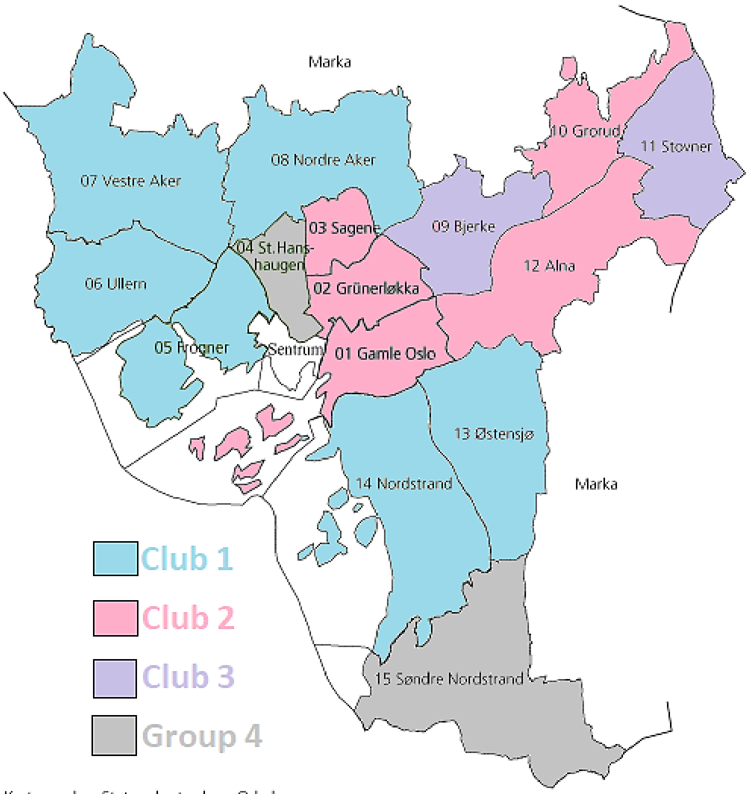

Figure 5.1: Map of Oslo, the city districts and the convergence (divergence) clubs (groups) formed by using total housing prices.

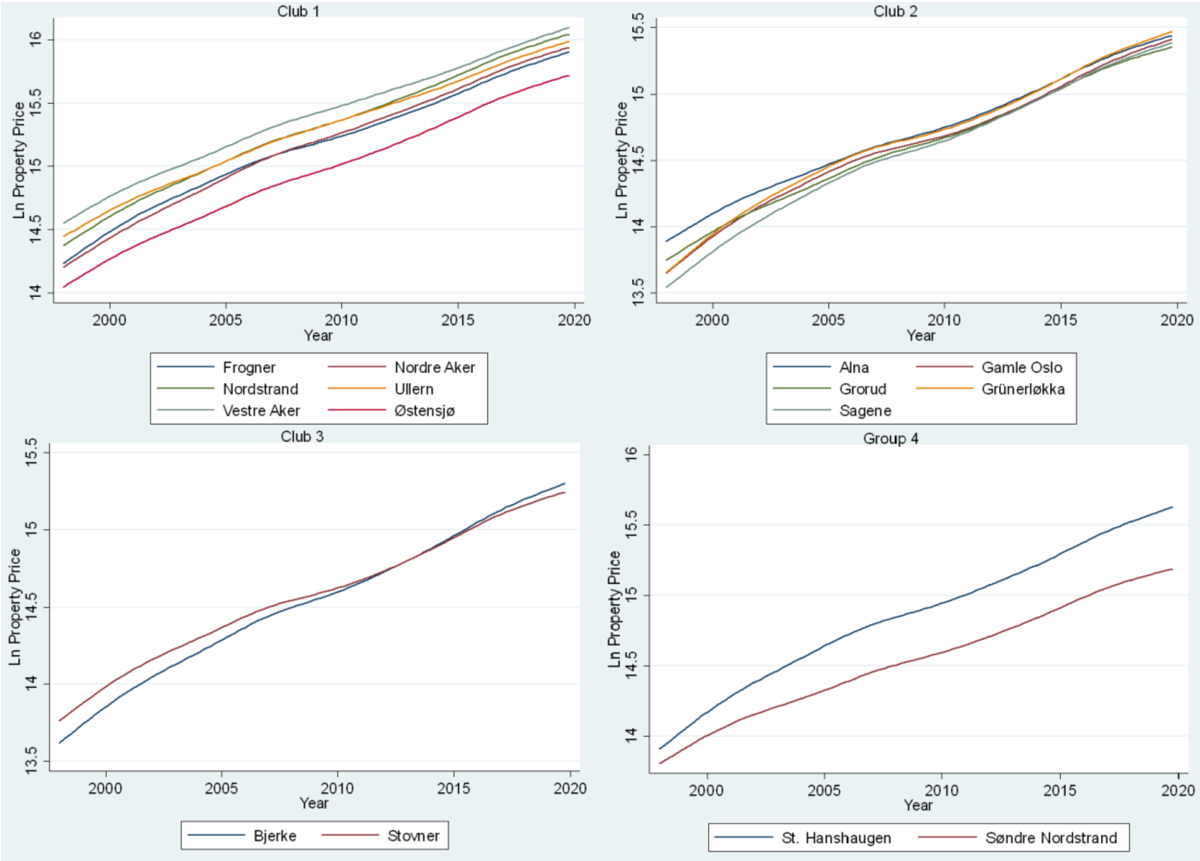

Figure 5.2: Graphs illustrating the clustering and divergence by showing the price appreciation for each district over time. Cyclical component filtered out.[18]

The graphs illustrated by Figure 5.2 shows how the housing prices in Club 1, Club 2 and Club 3 converges over time, while Group 4 clearly diverges. However, Club 3 seems to show a small amount of divergence after 2015. This can indicate that the converging pattern may turn around if given more time.

[18] See Figure A5.5 for the same graph with normal property prices.

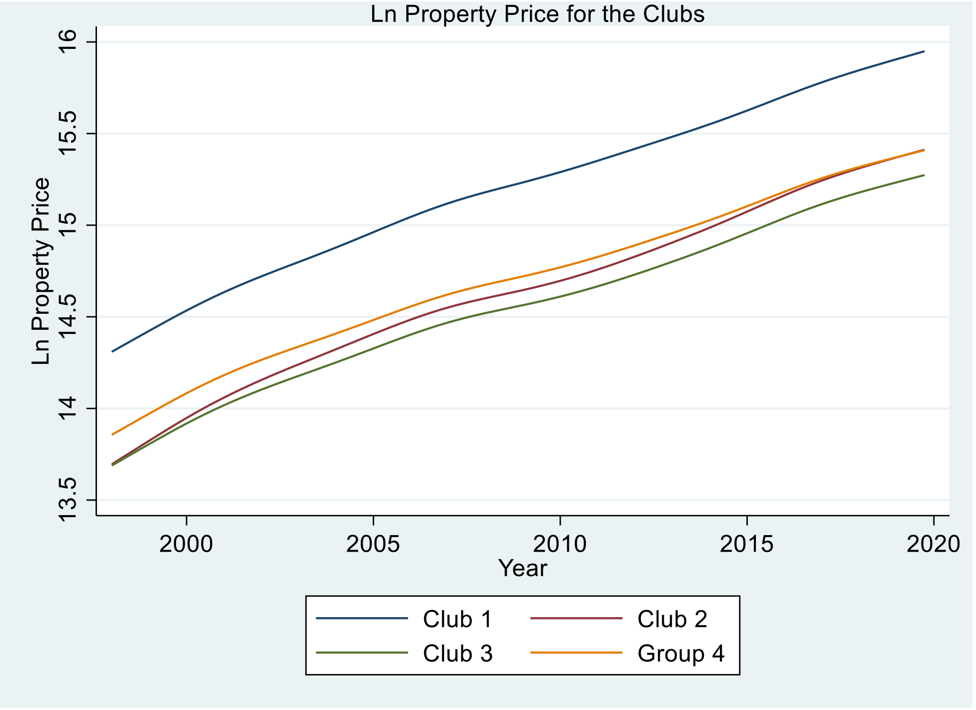

Figure 5.3: Graph showing the price appreciation measured in the natural logarithm for the respective convergence and divergence clubs. Cyclical component filtered out.[19]

In Figure 5.3, we see that Club 1 at the beginning of this time period (1998q1) had housing prices that were significantly higher than the other clubs. The position of Club 1 relative to the other clubs has not changed, but the price appreciation for this club has been way higher than the other groups. Club 2 have caught up with Group 4, and if the trend continues, Club 2 may pass Group 4. This means that the relative position between these subgroups have changed since the initial period. Club 3 and Group 4 seems to increase relatively similar to each other, but with a slight divergence.

[19] See Figure A5.6 for the same graph with normal property prices.

5.2 Overall convergence for housing prices per square meter with WRS

In the test for overall convergence Table 5.8 shows that we reject the null hypothesis of convergence in Oslo with a great margin, seeing as $t = -33.8249 < -1.65$.

Variable	Coefficient	Standard Error	T-Statistic
Ln Housing Price/sq.	-0.5591	0.0165	-33.8249*

*Table 5.8: Convergence test for overall convergence in Oslo measured in housing price per square meter. The number of individuals is 15. The number of time periods is 88. The first 26 periods are discarded before regression. * is denoted as a rejection of the null hypothesis of convergence at 5% significance level.*

Just like in section 5.1, the next step is to run convergence club tests to establish if there are any districts that that can form subgroups, where the housing price per square meter clusters towards a steady state. The results are provided in Table 5.9.

log(t)	Club1	Club2	Club3
Ln Housing Price/sq.	0.0421	0.2926	-0.1196
T-stat	2.2815	5.2179	-1.0041

*Table 5.9: The new convergence clubs after merging initial Club 2 and Club 3. * is denoted as the rejection of the null hypothesis of club merging at 5% significance level.*

Table 5.9 shows which city districts in Oslo the subgroups consists of. It is important to note that Club 3 barely passes the test for convergence at 5% significance level, with a t-level at -1.0041. This can indicate that there is only weak evidence for convergence. To determine if this is the case, we can look at this graphically, which will be done at a later stage in this section. Alna is the only district where the price per square meter does not converge with any other district.

Convergence club	***Districts***
Club 1	| Frogner | Gamle Oslo | Grünerløkka | Nordre Aker |
	| Sagene | St.Hanshaugen |
Club 2	| Bjerke | Grorud | Nordstrand | Ullern | Vestre Aker | Østensjø |
Club 3	| Stovner | Søndre Nordstrand |
No convergence	| Alna |

Table 5.10: Overview of the initial convergence clubs.

Finally, we will run a test for club merging to see if the housing prices per square meter between adjacent clubs clusters or not. The results are shown in Table 5.11. It is obvious that the null hypothesis is firmly rejected. This means that we keep the initial convergence clubs that are listed up in Table 5.10.

log(t)	Club1+2	Club2+3	Club3+G~4
Ln Housing Price/sq.	-0.2367	-0.6017	-1.2942
T-stat	-31.0259*	-28.6553*	-23.3932*

Table 5.11: Test for initial convergence club merging with adjacent clustering clubs. * is denoted as the rejection of the null hypothesis of club merging at 5% significance level.

If we compare Table 5.10 with Table 1.1 in Section 1, it is noteworthy to mention that Club 1 only consists of districts within outer and inner west and inner east. It only consists of districts in the center of Oslo or outer west of Oslo (Nordre Aker).

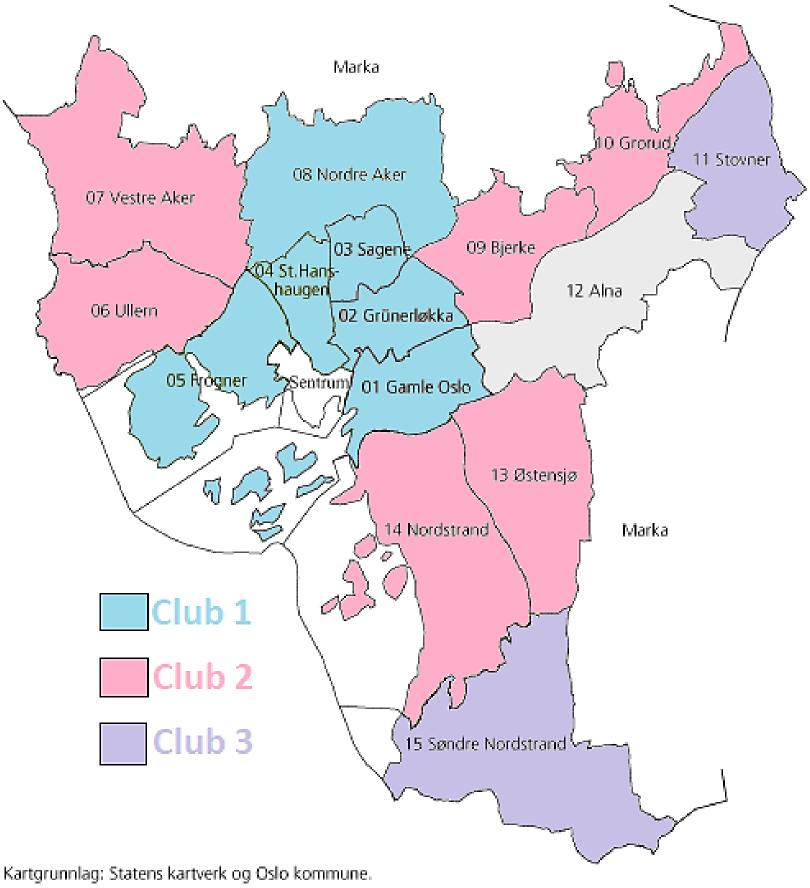

Figure 5.4: Map of Oslo, the city districts and the convergence (divergence) clubs (groups) formed by using housing prices per square meter.

When we look at the map of Oslo and its convergence clubs in Figure 5.4, we can see that all of the districts in Club 1 are also adjacent geographically to each other. In Club 2, the districts are more spread out in Oslo. However, there are three groups in this club that clusters with a neighbouring district. Club 3 does not consist of two neighbouring districts, but they still share very similar characteristics, i.e. ethnic composition, income and unemployment rate, despite of this. In fact, Figure A5.7, Figure A5.8 and Figure A5.9 in the Appendix shows that Club 3 have the most similar characteristics of all the clubs by a large margin.

[*Figure A5.7 found in the appendix*]
[*Figure A5.8 found in the appendix*]
[*Figure A5.9 found in the appendix*]

Neither Club 1 nor Club 2 have similar ethnic composition between the districts in the respective groups. The unemployment rate also differs quite a lot for said clustering clubs. Figure A5.10 shows the debt gearing in the clubs, and it shows that for both Club 2 and Club 3 the debt to yearly income ratio is quite similar between the districts. This is especially the case in more recent years.

[*Figure A5.10 found in the appendix*]

For Club 1 and Club 3, it seems like proximity and socioeconomic and demographic similarities are explainable factors. It is worth to note that the debt gearing has converged significantly in Club 3 between 2008 and 2018, which may somewhat explain the housing price convergence in this group. Club 2 is spread out and does not have many socioeconomic and demographic similarities, so there are probably some omitted explainable variables for convergence for the clustering club in question. Some important explainable factors may be proximity, school quality, general social infrastructure and consumption possibilities.

Figure 5.5 and Figure 5.6 illustrates graphically how the clubs are converging. The graphs below in Figure 5.5 does not show convergence at very fast rates. However, there is still a modest trend of clustering for each club, since the difference in housing prices per square meter reduces over time.

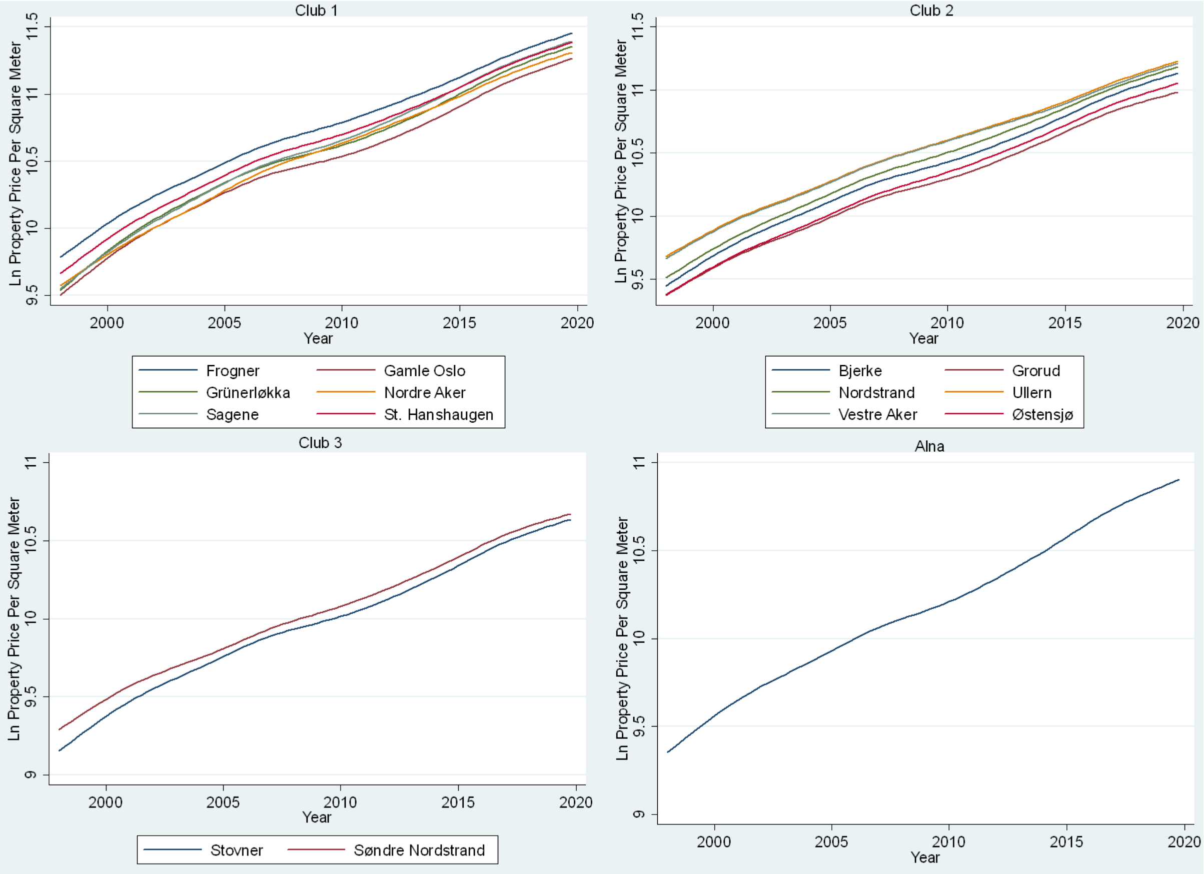

Figure 5.5: Graphs illustrating the club clustering by showing the price appreciation per square meter for each district over time. Cyclical component filtered out.[20]

In Figure 5.6 we can see that the convergence clubs and Alna have not changed their position in relative to the other subgroups. For example, Club 1 had the highest prices per square meter at the start of the period, which is still the case at the end of the time interval. The same can be said for the rest of the groups. Furthermore, the clustering groups clearly diverges from each other.

[20] See Figure A5.11 for the same graph with normal property prices.

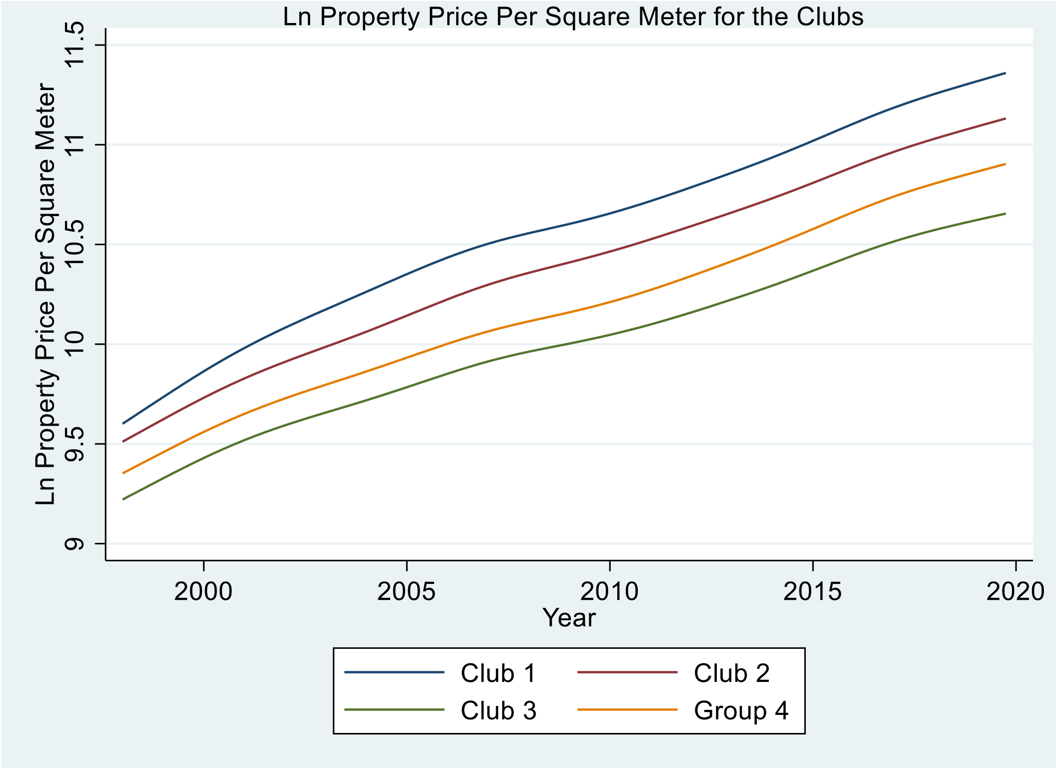

Figure 5.6: Graph showing the property price appreciation per square meter measured in the natural logarithm for the respective convergence clubs (and Alna). Cyclical component filtered out.[21]

5.3 Price determinants

This section consists of attempts to identify possible housing price determinants. Due to limitations of available data, there are only 24 observations for each convergence club. Also, many of the variables do not have quarterly data. Therefore, the results in the following section will only serve as indicators of price determinants.

5.3.1 Short term housing price changes

To identify price determinants, the following regression will be run for each convergence club identified in Oslo.

$$\begin{aligned}(18)\quad & DHP_{it} = \beta_0 + \beta_1 \ln(Inc_{it}) + \beta_2 IncRat_{it} + \beta_3 (IncChange_{it}) + \\ & \beta_4 ln\big(HP_{it}(-1)\big) + \beta_5 \ln(HPChange_{it}) + \beta_6 UnempRate_{it} + \beta_7 GDPChange_t + \\ & \beta_8 InterestRent_t + \beta_9 DebtGearing_{it} + \beta_9 NetMig_{it} + \beta_{10} Nor_{it} + \beta_{11} West_{it} + \\ & \beta_{12} EastEuro_{it} + \beta_{13} ResBuilt_{it} + \beta_{14} KM2_i + \beta_{15} Pop_{it} + \beta_{16} PopKM2_{it} + \varepsilon_{it}\end{aligned}$$

Where $DHPApr_{it}$ is the housing price appreciation in percentage, $IncRat_{it}$ is the ratio between the average income at district i at period t compared with the average income in

[21] See Figure A5.12 to see the same graph with normal property prices.

Oslo overall and $IncChange_{it}$ is the change in average income per capita. $HP_{it}(-1)$ are the housing prices in the period before t and $HPChange_{it}$ is the change in housing prices from $(t-2)$ and $(t-1)$, and these two variables represents the price expectations in period t. $UnempRate_{it}$ is the unemployment rate for each district, while $GDPChange_t$[22] is the change in GDP from $(t-1)$ to (t). $InterestRent_t$[23] represents the national real loan rent in Norway at period t. $DebtGearing_{it}$[24] is the ration between the average debt and the yearly income. $NetMig_{it}$ is the net migration. Nor_{it}, $West_{it}$ and $EastEuro_{it}$ is the share of people who are from Norway, Western countries[25] and East European countries not in the EU respectively. $ResBuilt_{it}$ is the number of new residences built, $KM2_i$ is the square kilometer in the district, Pop_{it} is the population and $PopKM2_{it}$ is the population per square kilometer.[26]

DHP_{it}	*Club 1*	*Club 2*	*Club 3*	*Group 4*
$\ln(Inc_{it})$	0.4416***	0.66149***	0.72931***	3.53303***
$IncRat_{it}$	-0.00315***	-0.01043***	-	-
$IncChange_{it}$	-0.00284	0.00621***	0.01364***	-0.03763***
$\ln(HP_{it}(-1))$	-0.31886***	-0.36371***	-0.48693***	-0.3432***
$\ln(HPChange_{it})$	0.08316	0.22619***	-0.01845	0.22125**
$UnempRate_{it}$	-0.01156	-0.00849	0.01***	0.04595**
$GDPChange_t$	0.0004	-0.00095	0.00233	0.0003
$InterestRent_t$	-0.00006	0.01998***	-0.01941***	0.0112***
$DebtGearing_{it}$	-0.11953	-0.03989	-0.21352*	-0.51385**
$NetMig_{it}$	-0.0008	-0.00184	0.02384***	-0.00781***
Nor_{it}	0.00254	-0.00121	-0.0089***	-0.03284***
$West_{it}$	0.00773	-0.00083	-0.01835	0.07466***
$EastEuro_{it}$	0.02752	0.00023	0.01232	-0.23198***
$ResBuilt_{it}$	2.33E-07	3.48E-08	-2.26E-06***	-3.32E-06***
$KM2_i$	0.02642*	0.00009	-0.09078*	-0.10425***
Pop_{it}	-0.00001*	0.00000131	-	-
$PopKM2_{it}$	0.00005	0.00001**	0.00002	-0.00021***
Constant	-0.78224	-2.21007***	-0.60941	-34.5387***
R-squared within	0.34286	0.43024	0.61436	0.6315
Overall r-squared	0.34521	0.43039	0.61436	0.63226
R-squared between	0.92129	0.80959	1	1

*Table 5.12: Random-effects robust GLS regression for each convergence club formed by using total housing prices. Each club has 24 observations per district. *** p < .01, ** p < .05, * p < .1.*

[22] Data received from Statistics Norway's Statbank, Table 09190 (GDP) and Table 01222 (Population).
[23] Data received from Statistics Norway (2020) "*Fakta om Norsk økonomi*".
[24] Data of debt received from Statistics Norway's Statbank, Table 05854. Income from footnote 6.
[25] West-Europe, USA, Canada, Australia and New Zealand
[26] Every variable denoted with an i is collected from Oslo municipality data bank.

In Table 5.12 we see the coefficients and statistical significance level for each convergence club. The first three variables indicate that income is statistically significant for explaining the variance of housing price appreciation. The change from $(t-2)$ to $(t-1)$ has a positive correlation for Club 2 and Group 4. Aside from economic status and housing price expectations, the statistically significant variables vary quite a lot between the clubs.

We can see that the independent variables explain a varying, but significant, amount of the variation of housing price appreciation by looking at the overall r-squared for each club. The r-squared within, which explains the variation within the same district, is very similar to the overall r-squared. We can also see that r-squared between is more than 0.8 for all the clubs. This means that the models explain a lot of the variation in price appreciation for each club.

5.3.2 Short term housing price changes per square meter

The same process will be used in this section to find price determinants for housing prices per square meter. The regression will be very similar to the one used in the previous section.

(19) $$DHPsq_{it} = \beta_0 + \beta_1 ln(Inc_{it}) + \beta_2 ln\big(HPsq_{it}(-1)\big) + \beta_3 \ln(HPsqChange_{it}) + \beta_4 UnempRate_{it} + \beta_5 GDPChange_{it} + \beta_6 InterestRent_{it} + \beta_7 NetMig_{it} + \beta_8 Nor_{it} + \beta_9 West_{it} + \beta_{10} EastEuro_{it} + \beta_{11} ResBuilt_{it} + \beta_{12} KM2_i + \beta_{13} Pop_{it} + \beta_{14} PopKM2_{it} + \varepsilon_{it}$$

The only variables that have been changed are $DHPsq_{it}$, $\ln(HPsq_{it}(-1))$ and $\ln(HPsqChange_{it})$, and they are the same as in Eq. 18, but for housing price per square meter. The results from the regressions can be found in Table 5.13.

$DHPsq_{it}$	*Club 1*	*Club 2*	*Club 3*
$ln(Inc_{it})$	0.53487***	0.44593***	1.03352
$IncRat_{it}$	-0.00515**	-0.00453***	-
$IncChange_{it}$	0.00093	-0.00359	0.00336
$ln(HPsq_{it}(-1))$	-0.3738***	-0.28675***	-0.50356***
$ln(HPsqChange_{it})$	0.32005***	0.02018	0.09614**
$UnempRate_{it}$	-0.03397	-0.02812**	0.07467**
$GDPChange_{t}$	-0.00039	0.00181*	-0.0016
$InterestRent_{t}$	0.00261	0.00035	-0.00446
$DebtGearing_{it}$	0.04918	-0.18282*	0.50421
$NetMig_{it}$	-0.00003	0.00607	-0.00502
Nor_{it}	-0.00237	-0.00475*	-0.04995
$West_{it}$	0.00737	0.03678	0.09418***
$EastEuro_{it}$	-0.01781	0.0024	-0.15311
$ResBuilt_{it}$	2.50E-0.8	4.98E-08	-1.42E-06
$KM2_{i}$	0.00185	-0.01208	-0.03019
Pop_{it}	1.32E-06	0.00001	-
$PopKM2_{it}$	0.00001	-0.0001	-0.00013
$Constant$	-2.33394*	-1.37383	-5.53564
R-squared within	0.38549	0.31837	0.63061
Overall r-squared	0.38604	0.32107	0.6321
R-squared between	0.80789	0.9362	1

*Table 5.13: Random effects robust GLS regression for each convergence club formed by using housing prices per square meter. Each club has 24 observations for each district. *** p < .01, ** p < .05, * p < .1.*

Club 3 is the only clustering group where income variables have no statistical significance. The price level at the previous period is negatively correlated with housing price appreciation for all the clubs. Price appreciation from $(t-2)$ to $(t-1)$ is positively correlated with price appreciation for Club 1 and Club 2. Unemployment rate is negatively correlated with price appreciation for Club 2, while the correlation is positive for Club 3. Overall, expectations and socioeconomic status seems to explain most of the variance in price appreciation. From the overall r-squared and r-squared within, we see that the explanatory power is over 0.3 for all the clubs. All clubs have an r-squared between that are above 0.8, which indicates that the models explains a lot of the variance between the districts in the respective clubs.

5.4 Comparing results from WRS with average prices

This section will use the index created by using average housing prices in Oslo. The same process will be used as in section 5.1 and 5.2, and therefore only the final results will be shown Table 5.14.

log(t)	Club1	Club2
Ln Housing Price	0.0216	-0.2999
T-stat	2.3979	-1.3662

Table 5.14: Test for final convergence club merging with adjacent clustering clubs at 5% significance level for total property prices.

Convergence club	Districts
Club 1	| BJERKE | FROGNER | GAMLE OSLO | GRÜNERLØKKA | NORDRE AKER | | NORDSTRAND | SAGENE | ST. HANSHAUGEN | ULLERN | VESTRE AKER | | ØSTENSJØ |
Club 2	| ALNA | GRORUD | STOVNER | SØNDRE NORDSTRAND |

Table 5.15: Overview of the convergence clubs as found by using an average price index. These are formed by using the total housing prices.

Table 5.15 shows the convergence clubs when not taking heterogeneity is not controlled for. As is immediately apparent, the results, and thus the groups, are quite different than the ones formed by using the WRS-index made by Case & Shiller.

In Table 5.16, the convergence test results by using housing price per square meter and an average housing price index is summarized.

Log(t)	Club1	Club2	Club3	Group4
Ln Housing Price/sq	0.1907	0.8734	0.5703	-0.3963
T-stat	11.8012	20.6622	1.8061	-7.0903

Table 5.16: Test for final formation of convergence clubs at 5% significance level for housing prices per square meter. * is denoted as the rejection of the null hypothesis of club merging at 5% significance level.

Convergence clubs	Districts
Club 1	| FROGNER | GAMLE OSLO | GRÜNERLØKKA | NORDRE AKER | SAGENE | | ST. HANSHAUGEN |
Club 2	| BJERKE | NORDSTRAND | VESTRE AKER | ØSTENSJØ |
Club 3	| STOVNER | SØNDRE NORDSTRAND |
Group 4	| ALNA | GRORUD | ULLERN |

Table 5.17: Overview of the convergence clubs as found by using an average price index. These are formed by using the housing prices per square meter.

The formed convergence clubs when using housing prices per square meter are quite similar when using an average and the WRS-index. Club 1 and Club 3 are completely the same for both indexes. When using the WRS-index, Club 2 consists of Bjerke, Nordstrand, Vestre Aker and Østensjø as well as Grorud and Ullern. This means that Club 2 and Group 4 differs somewhat between the two indexes. Table 5.15 and Table 5.17 shows that the results from the convergence test can differ somewhat depending on the index used. Since the average housing price index does not account for heterogeneity, Table 5.15 and Table 5.17 shows the results when not adjusting for changes in the housing market.

6. Conclusion

The convergence tests done in this thesis shows that there is no overall convergence in Oslo, and that subgroups that clusteres towards a steady state has been identified. For total housing prices, there are three clustering clubs, while one group diverges. The respective clubs have mostly kept their price position relative to the other subgroups. Club 2 is the exception here, who seems to have had a faster appreciation rate than the rest of the groups. It seems like proximity and socioeconomic similarities can explain some of the price convergence.

Three convergence clubs were identified when measured in prices per square meter. Alna was not included in any of these groups. When comparing the identified convergence clubs by using WRS and average housing prices, the results are quite similar. For instance, Club 1 and Club 3 are completely identical. Proximity to neighbouring districts and socioeconomic similarities seem to explain some of the variance between the districts in the convergence clubs. There seems to be a geographic component in the formation of the clustering groups, especially for Club 1. This can indicate that there are similar geographic amenities within the clubs that impacts the price development. Club 2 is split up into three neighbouring districts, who are spread out in different parts of Oslo. Since the districts in Club 2 and Club 3 are quite split up, this can suggest that the geographic amenities can change over time.

The club formation, when using prices per square meter, in Oslo suggests that the development in the housing market perpetuates the growing ethnic and socioeconomic segregation. As shown in Figure 5.6, the convergence clubs in Oslo clearly diverges from each other. If this pattern of divergence continues, the unfortunate consequence is further segregation and geographic inequality.

The random effects robust GLS-models in section 5.3 strongly indicate that socioeconomic status as well as expectations of housing prices at period t explains a lot of the variance in the prices and the appreciation rate. These attempts of identifying the price determinants were heavily limited by the data, available time and the scope of the thesis. Considering the negative effects of having segregated communities, a recommendation for future research is to conduct a more thorough analysis of what determines the formation of the convergence and divergence clubs. Such a study would be a great contribution to the literature on the housing market in Oslo.

7. References

Abbott, A. & Vita, G. (2013). "*Testing for long-run convergence across regional housing prices in the UK: A pairwise approach*". Applied Economics. 45. 1227-1238. 10.1080/00036846.2011.613800. (Read 08.10.2020)

Abraham, J. & Schauman, W. S. (1991). "*New Evidence on Home Prices from Freddie Mac Repeat Sales*". *AREUEA Journal*, vol. 19, no. 3, pp. 333-52. (Read 16.10.2020)

Alsvik, B. (2009). "Kristianiakrakket 1899". Oslo kommune, Byarkivet. Received from: https://www.oslo.kommune.no/OBA/tobias/tobiasartikler/Tob2009-2_5.htm
(Read 01.12.2020)

Apergis, N. & Simo-Kengne, B. & Gupta, R. (2015). "*Convergence in provincial-level South African housing prices: Evidence from the club convergence and clustering procedure*". Review of Urban & Regional Development Studies. 27. 10.1111/rurd.12029.
(Read 08.10.2020)

Awaworyi Churchill, S. & Inekwe, J. & Ivanovski, K. (2018). "*Housing price convergence: Evidence from Australian cities*". Economics Letters. 170. 88-90. 10.1016/j.econlet.2018.06.004. (Read 15.09.2020)

Bailey, M. J., Muth, R. F. & Nourse, H. O. (1963). "*A Regression Method for Real Estate Price Index Construction*". Journal of American Statistical Association, 58, 933-942. http://dx.doi.org/10.1080/01621459.1963.10480679 (Read 13.11.2020)

Barstad, H. (2016). "*Et godt sted å bo*". Oslo kommune, Byarkiv. Received from: https://www.oslo.kommune.no/OBA/tobias/tobiasartikler/pdf_arkiv/tobias2016.pdf
(Read 01.12.2020)

Baug, P. & Dyvi, Y. (2008). "*MODAG – en makroøkonomisk modell for norsk økonomi*". Statistisk Sentralbyrå. Received from: https://www.ssb.no/nasjonalregnskap-og-konjunkturer/artikler-og-publikasjoner/modag-en-makrookonomisk-modell-for-norsk-okonomi. (Read 20.10.2020)

Carlino, G., Mills, L. (1993). "*Are US incomes converging? A time series analysis"*. Journal of Monetary Economics 32, 335–346. (Read 02.10.2020)

Case, K. E., & Shiller, R. J. (1989). "*The efficiency of the market of single-family homes*". The American Economic Review, 79(1), 125–137. (Read 13.11.2020)

Case, B., Pollakowski, H. O. & Wachter, S. M. (1997). "*Frequency of Transaction and Housing Price Modeling*". Journal of Real Estate Finance and Economics, vol. 14, pp. 173-87. (Read 17.10.2020)

Case, B. & Quigley, J. M. (1991). "*The Dynamics of Real Estate Prices*". The Review of Economics and Statistics, vol. 73, no. 1, February, pp. 50-58. (Read 17.10.2020)

Crone, T. M. & Voith, R. P. (1992). "*Estimating housing price appreciation: A comparison of methods*". Journal of Housing Economics, Elsevier, vol. 2(4), pages 324-338, December. (Read 18.10.2020)

Clapp, J.M., Giaccotto, C. & Tirtiroglu, D. (1991). "*Housing Price Indices Based on All Transactions Compared to Repeat Subsamples*". Real Estate Economics, 19: 270-285. doi:10.1111/1540-6229.00553 (Read 18.10.2020)

Cook, S. (2003). "*The Convergence of Regional Housing Prices in the UK*". Urban Studies. 40. 2285-2294. 10.1080/0042098032000123295. (Read 08.10.2020)

Cooper C., Orford S., Webster C., et al. (2013). "*Exploring the Ripple Effect and Spatial Volatility in Housing Prices in England and Wales, Regressing Interaction Domain Cross-Correlations Against Reactive Statistics*". Environment and Planning B, Planning and Design, Vol. 40, No. 5, pp. 763–782, doi: 10.1068/b37062. (Read 06.10.2020)

Du, K. (2018). "*Econometric Convergence Test and Club Clustering Using Stata*". The Stata Journal: Promoting communications on statistics and Stata. 17. 882-900. 10.1177/1536867X1801700407. (Read 03.09.2020)

Dybendal, K. E. & Skiri, H. (2005). "*Klare geografiske forskjeller i levealder mellom bydeler i Oslo".* Statistisk Sentralbyrå. Received from: https://ssb.brage.unit.no/ssb-xmlui/handle/11250/179202 (Read 28.08.2020)

Eiendom Norge (n. d.). "Om statistikken". Received from: https://eiendomnorge.no/boligprisstatistikk/om-statistikken/ (Read 15.12.2020)

Eiendomsverdi (n. d.) "*Eiendomsverdi*". Received from: https://eiendomsverdi.no/ (Read 29.10.2020)

Gabrieli, T. & Panagiotidis, T. & Xu, Y. (2019). "*Pair-wise convergence of intra-city housing prices in Beijing*". Journal of Economic Asymmetries. 20. 10.1016/j.jeca.2019.e00131. (Read 08.10.2020)

Ganioğlu, A. & Seven, Ü. (2019). "*Convergence in Housing Prices: Cross-Regional Evidence for Turkey*". (Read 08.10.2020)

Grytten, O. (2009). "Boligkrisen - Norges plass i det globale boligkrakket". *Civita-Notat nr 3*, Oslo. Received from: https://www.civita.no/publikasjon/boligkrisen-norges-plass-980 (Read 02.12.2020)

Gupta R. & Miller S. M. (2012) "*Ripple effects and forecasting home prices in Los Angeles, Las Vegas and Phoenix*". The Annals of Regional Science, Vol. 48, No. 3, pp. 763–782, doi: 10.1007/s00168-010-0416-2. (Read 10.10.2020)

Harding, J. P., Rosenthal, S. S. & Sirmans, C. F. (2007). "*Depreciation of Housing Capital, Maintenance, and Housing Price Inflation: Estimates from a Repeat Sales Model*" *Journal of Urban Economics*, forthcoming. (Read 16.10.2020)

Hager-Thoresen, F., Murray, S. M. & Sørgjerd, C. (2020). "*Enorme forskjeller innad i Oslo: Her øker boligprisene mest og minst*". Aftenposten. Received from: https://www.aftenposten.no/oslo/i/aPOM4A/se-oversikten-saa-store-er-prisforskjellene-innadioslo?fbclid=IwAR3Ov6KbDEFQCqguExpfMK38gM8BxAgzt3rFUNgq9lKoDp7C7Why20zOnjo (Read 01.12.2020)

Hernæs, Ø., Markussen, S. & Røed, K. (2020). "*Økende inntektssegregering i norske byregioner*". Tidsskrift for boligforskning. Received from: https://www.idunn.no/tidsskrift_for_boligforskning/2020/02/oekende_inntektssegregering_i_norske_byregioner (Read 01.12.2020)

Hodrick, R. J. & Prescott, E. C. (1997). "*Postwar U.S. business cycles: An empirical investigation*". Journal of Money, Credit, and Banking 29: 1–16. (Read 14.10.2020)

Kelly, M. (2000). "*Inequality And Crime*". The Review of Economics and Statistics. 82. 530-539. 10.1162/003465300559028. (Read 26.08.2020)

Kim, Y. S. & Rous, J. J. (2012). "*Housing price convergence: Evidence from US state and metropolitan area panels*," Journal of Housing Economics, Elsevier, vol. 21(2), pages 169-186. (Read 08.10.2020)

Kirkeberg, M. I. (2003). "*Oslo - flest fattige og størst ulikhet*". Statistisk Sentralbyrå. Received from: https://ssb.brage.unit.no/ssb-xmlui/handle/11250/179523 (Read 28.08.2020)

Kongsrud, P. (2000). "*Forstår vi prisdannelsen i boligmarkedet?*". Finansdepartementet. Received from: https://www.regjeringen.no/no/aktuelt/Forstar-viprisdannelsen-i-boligmarkedet/id423430/ (Read 18.10.2020)

Krakstad, S. & Oust, A. (2015). "*Are housing prices in the Norwegian capital too high?*". International Journal of Housing Markets and Analysis. 8. 152-168. 10.1108/IJHMA-08-2014-0034. (Read 25.08.2020)

Krogsveen (2020). "*Prisutvikling fra Oslo*". Received from: https://www.krogsveen.no/prisstatistikk/oslo (Read 25.08.2020)

Kuketayev, A. (2013). "*The convergence of regional housing prices in the USA in the context of the stress testing of financial institutions*". (Read 08.10.2020)

Kutluay, M., Vale, P. H. & Yildiz, S. (2015). "*Er det boligboble i Norge?*" Magma – Econas Tidsskrift for økonomi og ledelse. Received from: https://www.magma.no/er-det-boligboble-i-norge (Read 02.12.2020)

Lilleby, J. (2018). "*Kristiania-krakket: Da boligboblen sprakk*". E24. Received from: https://e24.no/norsk-oekonomi/i/Adb91n/kristiania-krakket-da-boligboblen-sprakk (Read 20.11.2020)

Liri, R. & Zhang, X. & Li, X. & Dong, J. (2015). "*Heterogeneous convergence of regional housing prices and the complexity in China*". Zbornik Radova Ekonomskog Fakulteta u Rijeci / Proceedings of Rijeka School of Economics. 33. 325-348. 10.18045/zbefri.2015.2.325. (Read 06.10.2020)

MacDonald, R. & Taylor, M. (1993). "*Regional housing prices in Britain: Long-run relationships and short-run dynamics*". *Scottish Journal of Political Economy*, *40*, 43–55. (Read 08.10.2020)

Meen, G. (1999). "*Regional housing prices and the ripple effect: a new interpretation*". Housing Studies. 14, pp. 733-753. (Read 09.10.2020)

Meese, R. A. & Wallace, N. E. (1997). "*The Construction of Residential Housing Price Indices: A Comparison of Repeat-Sales, Hedonic- Regression, and Hybrid Approaches*". Journal of Real Estate Finance and Economics, vol. 14, pp. 51-73. (Read 18.10.2020)

Montagnoli, A. & Nagayasu, J. (2015). "*UK housing price convergence clubs and spillovers"*. Journal of Housing Economics. 30. 50-58. 10.1016/j.jhe.2015.10.003. (Read 07.10.2020)

Montañés bernal, A. & Olmos, L. (2013). "*Convergence in US housing prices*". Economics Letters. 121. 152-155. 10.1016/j.econlet.2013.07.021. (Read 08.10.2020)

Natsvaladze, M. & Beraia, N. (2018). "*Real estate Economics*". 10.13140/RG.2.2.13631.41124. (Read 14.10.2020)

Ng, S. & Perron, P. (2001) “*Lag Length Selection and the Construction of Unit Root Test with Good Size and Power*”. Econometrica, Vol. 69, No. 6, pp. 1519–1554; doi: 10.1111/1468-0262.00256. (Read 02.10.2020)

Norgeshistorie (2019). “*Christies gate*”. Universitetet i Oslo. Received from: https://www.norgeshistorie.no/studere-fortid/studenttekster/2019/S19_14_christies-gate.-kristianiakrakkets-konsekvenser.html (Read 20.11.2020)

Norsk takst (2015). “*Takstbransjens retningslinjer ved arealmåling -2014*”. Eiendomsforlaget AS. Received from: https://www.norsktakst.no/media/1158/takstbransjens-retn-linjer-ved-arealmaaling-2014-2-utgave-9215-1-.pdf (Read 15.09.2020)

OBOS (n. d.). “*En av Nordens Største boligbyggere*”. Received from: https://nye.obos.no/dette-er-obos/historien-om-obos/ (02.12.2020)

Oslo Kommune (n. d.). “*Geografiske inndelinger*”. Received from: https://www.oslo.kommune.no/statistikk/geografiske-inndelinger/ (Read 09.12.2020)

Oslo Kommune Statistikkbanken (n. d.). “Oslo kommunes statistikkbank”. Received from: http://statistikkbanken.oslo.kommune.no/webview/ (Read 04.12.2020)

Peek, J. & Wilcox, J. A. (1991). “*The Measurement and Determinants of Single-Family Housing Prices*” *ARUEA Journal*, vol. 19, no. 3, pp. 353-82. (Read 16.10.2020)

Phillips, P. C. B. & Sul, D. (2007). “*Transition modelling and econometric convergence tests”.* Econometrica Vol. 75, p. 1771-1855. (Read 05.09.2020)

Phillips, P. C. B. & Sul, D. (2009). “*Economic transition and growth. Journal of Applied Econometrics*”. 24, (7), p. 1153-1185. Research Collection School of Economics. (Read 05.09.2020)

Prognosesenteret (2014). "*Om boligmarked og boligfinansiering*". Finans Norge. Received from: https://www.finansnorge.no/contentassets/f2ad4fbbfe38459f80c21970d24c1a63/boligmarked-og-boligfinansiering.pdf (Read 14.11.2020)

Rappaport, J. (2007). "*A Guide to Aggregate Housing Price Measures*". Economic Review. 92. 41-71. (Read 02.09.2020)

Schnurbus, J., Haupt, H. & Meier, V. (2017). "*Economic Transition and Growth: A Replication*". Journal of Applied Econometrics. 32. 1039-1042. 10.1002/jae.2544. (Read 14.10.2020)

SpareBank (n.d.). Eiendomsverdi: "*Automated Valuation Company for Residential Real Estate*". Retrieved from: https://spabol.sparebank1.no/articles/eiendomsverdi-automated-valuation-company-for-residential-real-estate (Read 29.10.2020)

Statistikkbanken (2018). "*Skatt for personer 05854: Hovedposter fra skatteoppgjøret for bosatte personer 17 år og eldre, etter alder. Gjennomsnitt og median (kr) (K) (B) 1999 – 2018*". Statistisk sentralbyrå. Received from: https://www.ssb.no/statbank/table/05854/ (Read 04.12.2020)

Statistikkbanken (2020). "*Befolkning 01222: Endringar i befolkninga i løpet av kvartalet, for kommunar, fylke og heile (K) 1997K4 – 2020K3*". Statistisk sentralbyår. Received from: https://www.ssb.no/statbank/table/01222/tableViewLayout1/ (Read 18.10.2020)

Statistikkbanken (2020). "Nasjonalregnskap 09190: Makroøkonomiske hovedstørrelser. Ujuster og sesongjustert 1978K1 - 2020K3". Statistisk sentralbyrå. Received from: https://www.ssb.no/statbank/table/09190/ (Read 18.10.2020)

Statistisk sentralbyrå (2020). "*Fakta om Norsk økonomi*". Received from: https://www.ssb.no/nasjonalregnskap-og-konjunkturer/faktaside/norsk-okonomi (Read 19.10.2020)

Torsvik, R. M. (1999). “*Bankkrisen*”. Statistisk sentralbyrå. Received from: https://www.ssb.no/bank-og-finansmarked/artikler-og-publikasjoner/bankkrisen (Read 02.12.2020)

Tuv, N (2019). “*Disse kommunene har størst og minst inntektsulikhet*”. Statistisk sentralbyrå. Received from: https://www.ssb.no/inntekt-og-forbruk/artikler-og-publikasjoner/disse-kommunene-har-storst-og-minst-inntektsulikhet (Read 25.08.2020)

Wong, W., De Silva, A. (2015). “*Housing price convergence: a Melbourne case study*”. Proceedings from the 21st Pacific-Rim Real Estate Society Conference - 2015, Kuala Lumpur, Malaysia. (Read 08.10.2020)

Żelazowski, K. (2019). “*Price Convergence in the Regional Housing Markets in Poland, Real Estate Management and Valuation*”, 27(2), 44-52. doi: https://doi.org/10.2478/remav-2019-0014 (Read 08.10.2020)

Øia, T. (2007). “*Ung i Oslo Levekår og sosiale forskjeller”.* Norsk institutt for forskning om oppvekst, velferd og aldring. Received from: https://digib.no/wp-content/uploads/2020/01/153.pdf (Read 28.08.2020)

Øye, B.-E. (2019). “*Oslo på top – Tønsberg på bunn*”. Smartly. Received from: https://baerekraft.smartly.no/oslo-paa-topp--toensberg-paa-bunn (Read 19.10.2020)

8. Figures

Figure 1.1: Oslo Kommune (2017). "*Bydelskart for* Oslo". Retreived from: https://www.oslo.kommune.no/getfile.php/13206469-1490274697/Tjenester%20og%20tilbud/Politikk%20og%20administrasjon/Statistikk/Geografiske%20inndelinger/Oslo_Bydelskart_20170221_A3.pdf (Read 28.08.2020)

Figure 5.1: Dybendal, K. E. (2005). "*Klare geografiske forskjeller i levealder mellom bydeler i Oslo*". Statistisk sentralbyrå. Received from: https://www.ssb.no/befolkning/artikler-og-publikasjoner/klare-geografiske-forskjeller-i-levealder-mellom-bydeler-i-oslo
(Read 03.12.2020)

Figure 5.4: Dybendal, K. E. (2005). "*Klare geografiske forskjeller i levealder mellom bydeler i Oslo*". Statistisk sentralbyrå. Received from: https://www.ssb.no/befolkning/artikler-og-publikasjoner/klare-geografiske-forskjeller-i-levealder-mellom-bydeler-i-oslo
(Read 03.12.2020)

Figure A2.2: Prognosesenteret (2014). "*Om boligmarked og boligfinansiering*". Finans Norge. Retreived from: https://www.finansnorge.no/contentassets/f2ad4fbbfe38459f80c21970d24c1a63/boligmarked-og-boligfinansiering.pdf (Read 14.11.2020)

Figure A2.3: Rappaport, J. (2007). "*A Guide to Aggregate Housing Price Measures*". Economic Review. 92. 41-71. (Read 02.09.2020)

Appendix

Appendix - Figures

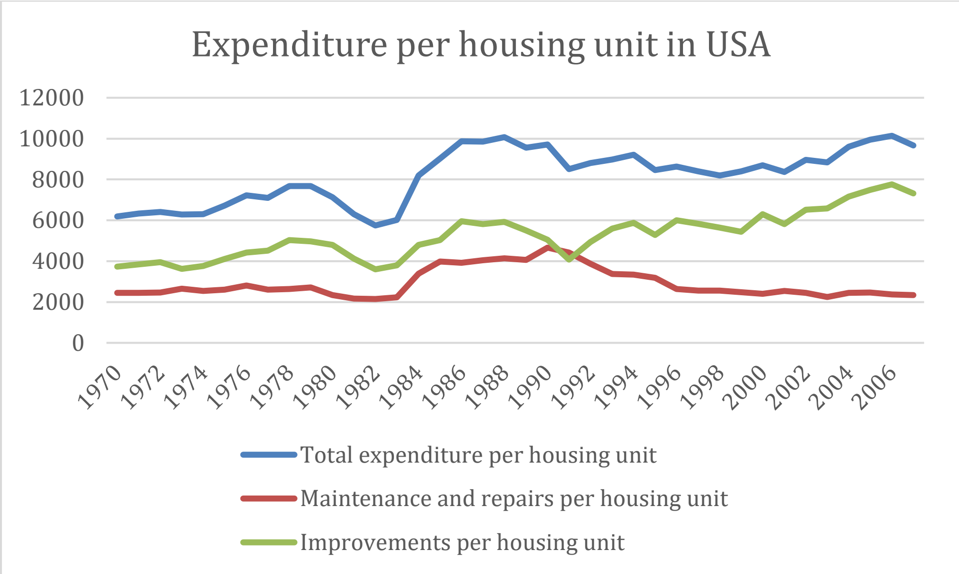

Figure A2.1: Graph showing expenditure per housing unit in the US in 2020 US dollars. Created the graphs based on data from the US Census Bureau.

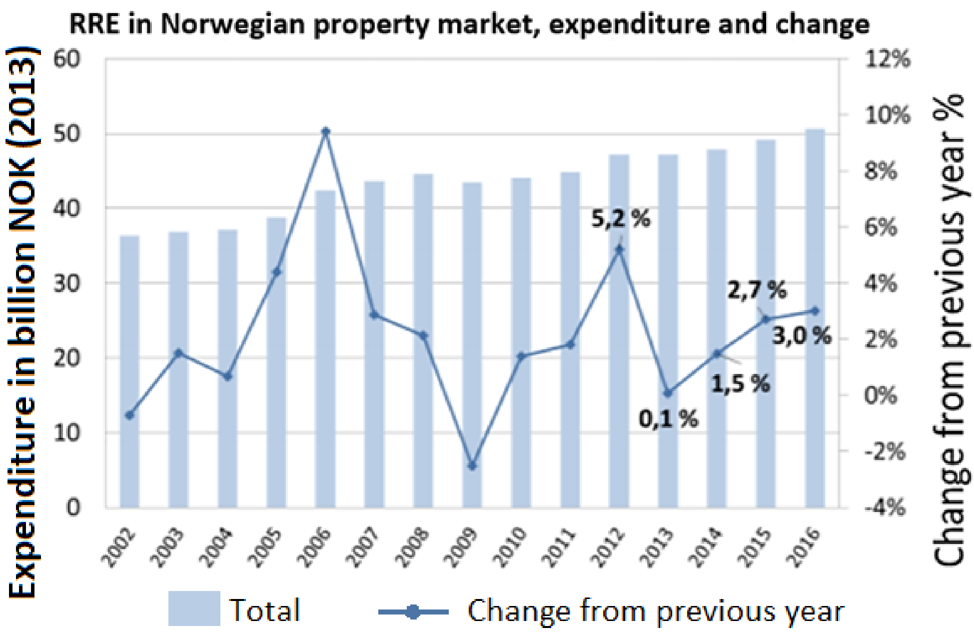

Figure A2.2: A graph showing the expenditure on renovation, rebuilding and extension (RRE) of existing properties. The left axis shows the total expenditure in billion NOK (2013-prices), while the right axis shows the change from the previous year.

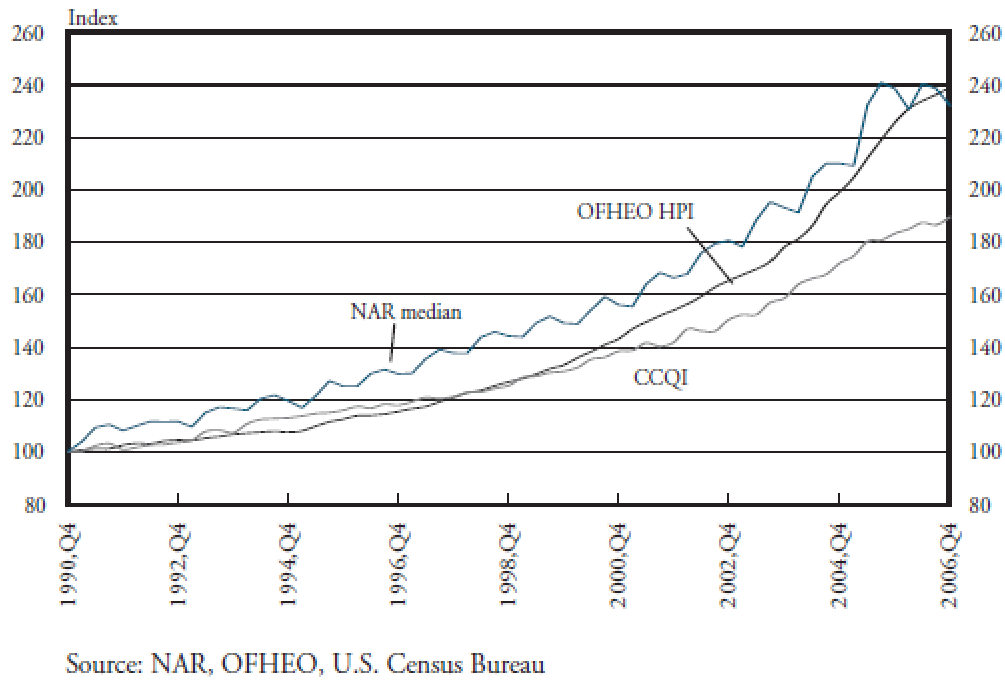

Figure A2.3: Graph showing the appreciation path for the average methodology (NAR median), the repeat sales methodology (OFHEO HPI) and the hedonic methodology (CCQI).

One unfortunate disadvantage with the graphs in Figure 2.3, is that CCQI only measures the aggregate average home prices for new homes, while the two others measures the average for existing homes.

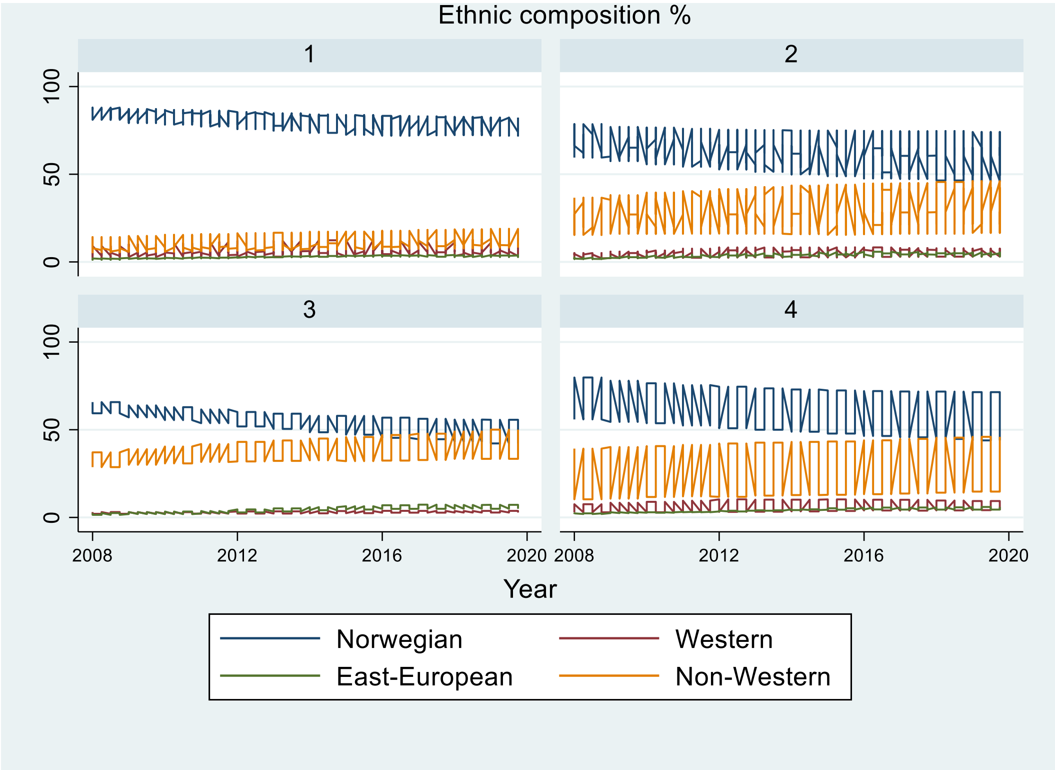

Figure A5.1: Ethnic composition for convergence clubs identified by using total property prices. Shows the share of the population for each district at period t*. The thicker the graphs are, the greater the difference.*

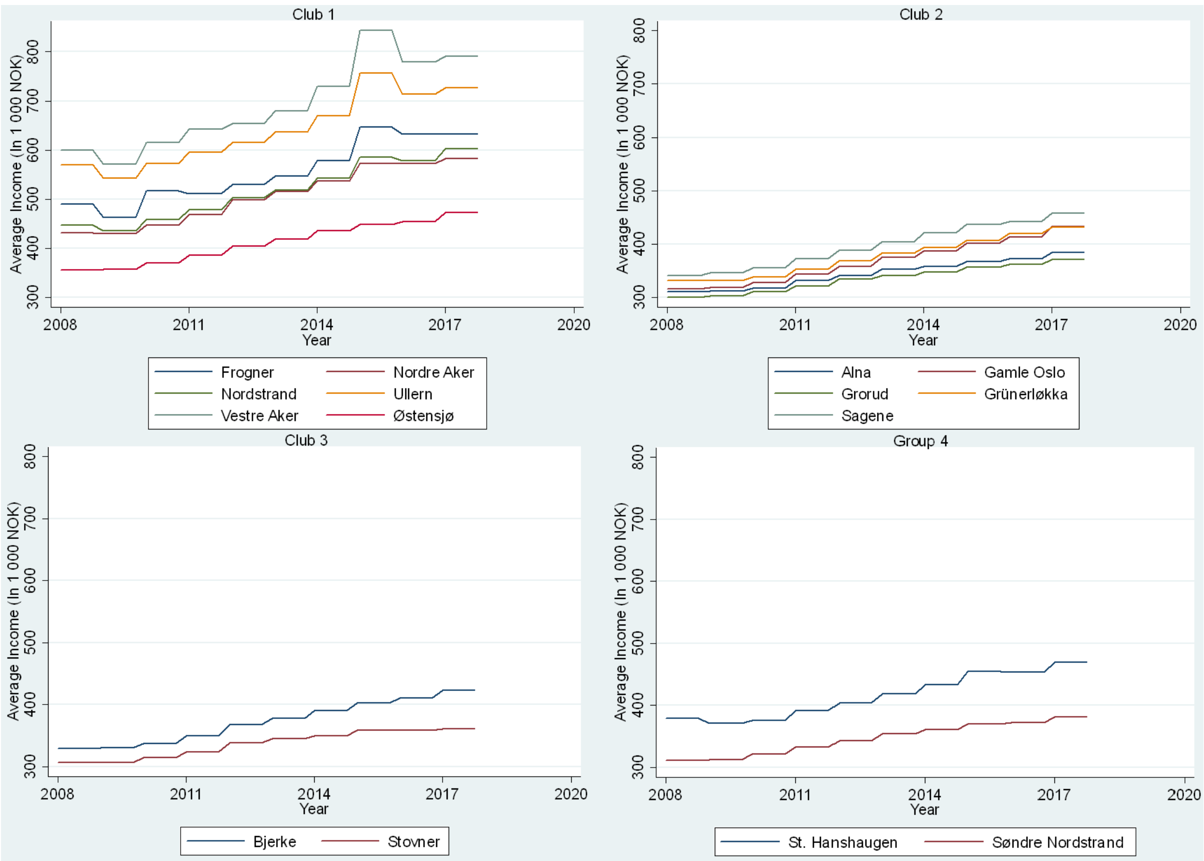

Figure A5.2: Level of income for the clubs identified by using total property prices. Shows the average income of the population for each district at period t*.*

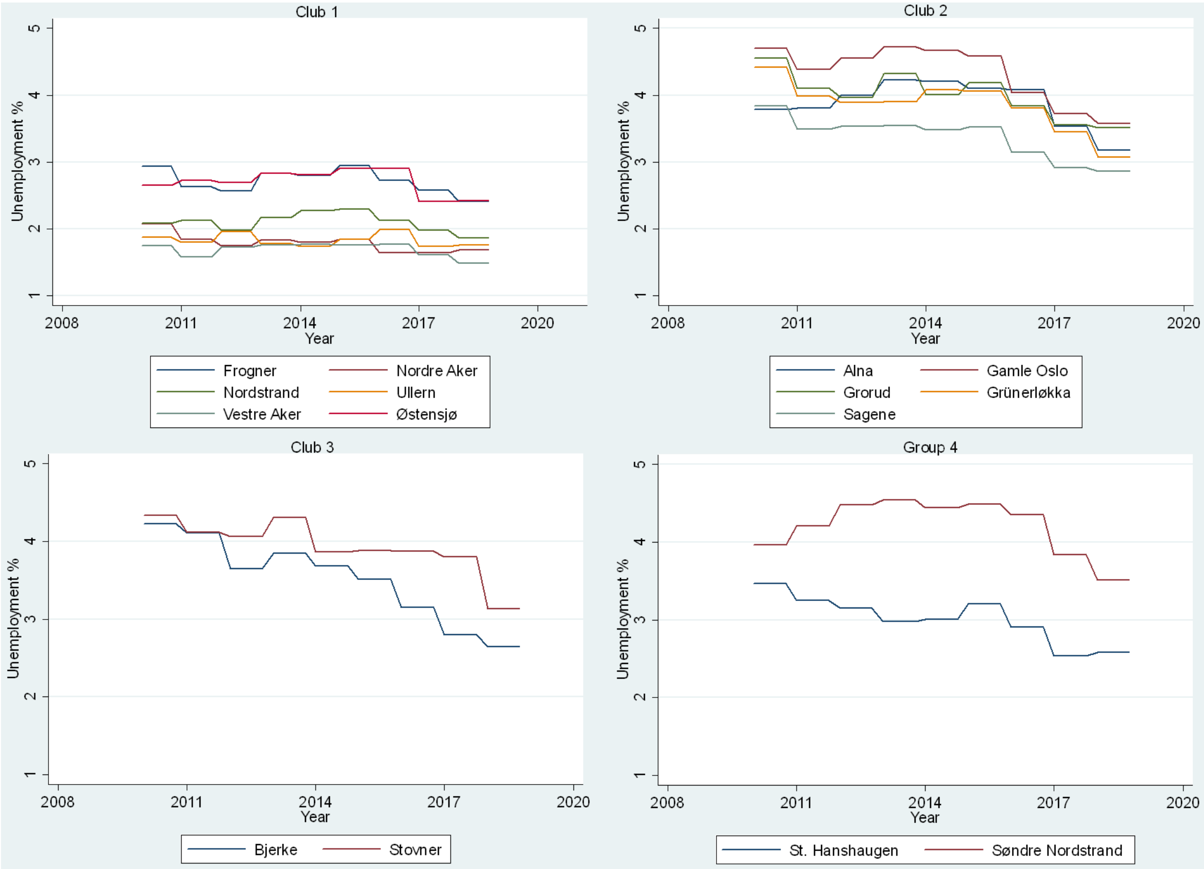

Figure A5.3: Unemployment rate for the clubs identified by using total property prices. Shows the unemployment rate for each district at period t.

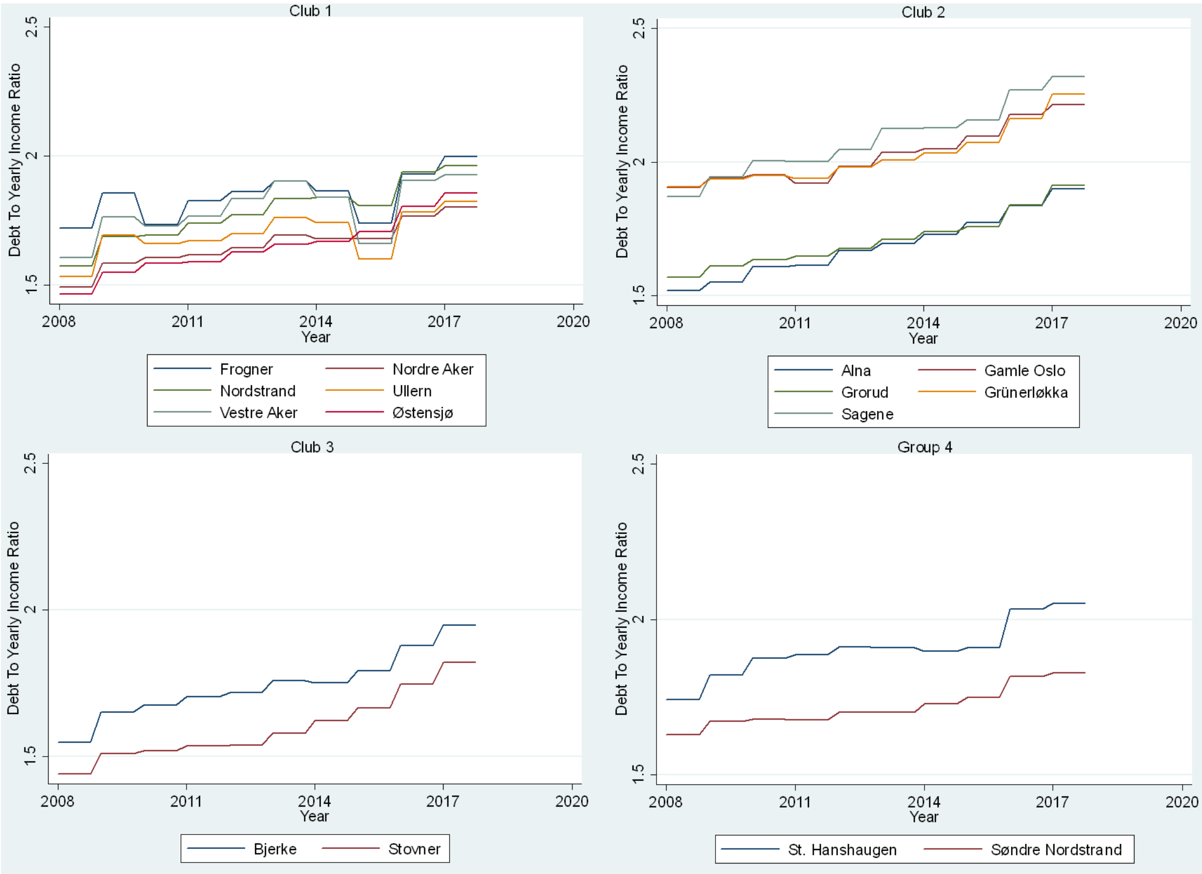

Figure A5.4: Debt to yearly income ratio in the respective convergence clubs identified by using total property prices. Found by dividing the total average debt with the average yearly income. Shows the debt gearing of the population for each district at period t.

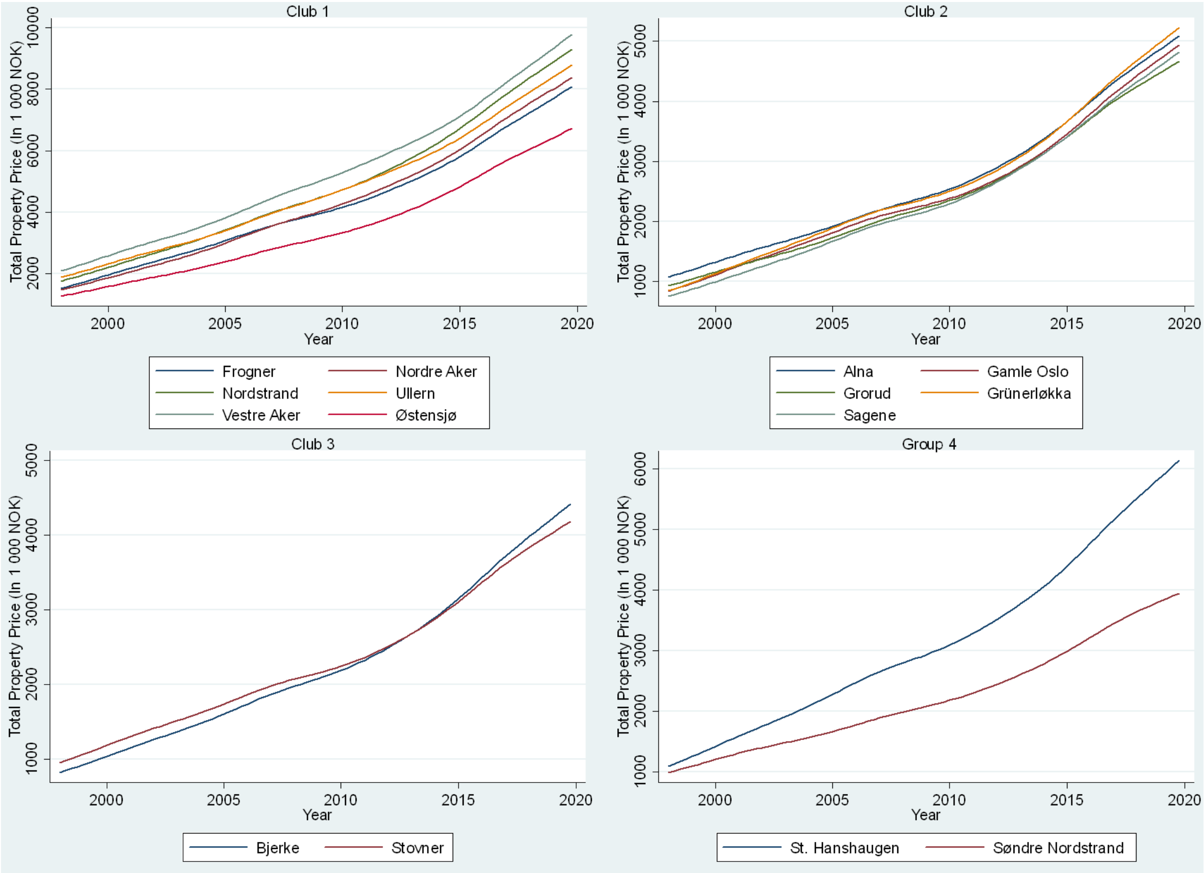

Figure A5.5: Graphs illustrating the total price appreciation in normal values (in 1 000 NOK) for all the clubs and their respective districts. Convergence clubs identified by using total property prices.

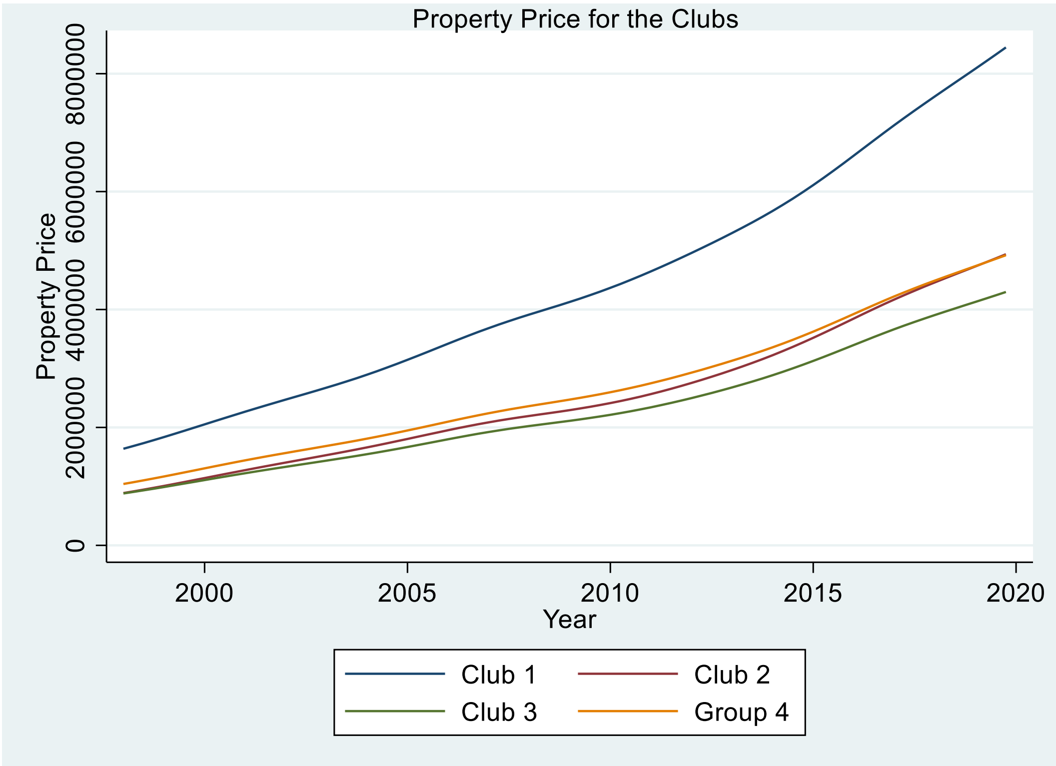

Figure A5.6: Graphs showing the price appreciation for the convergence clubs in normal values. Convergence clubs identified by using total property prices.

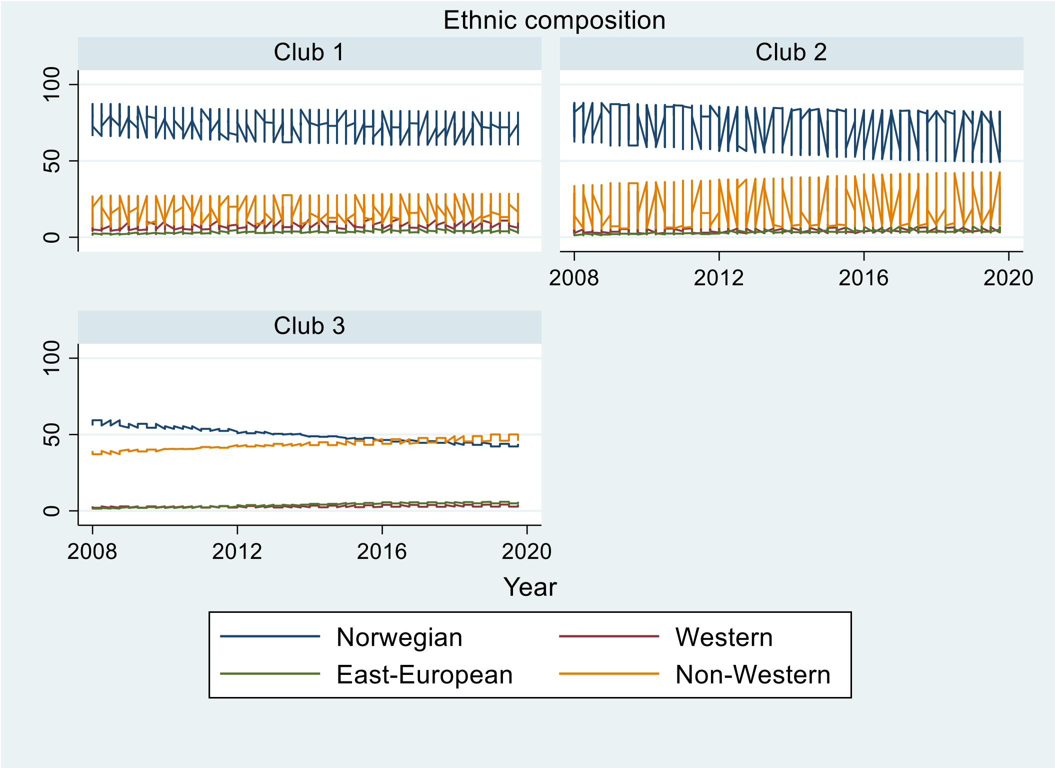

Figure A5.7: Ethnic composition for convergence clubs identified by using property prices per square meter. Shows the share of the population for each district at period t. The thicker the graphs, the greater the difference.

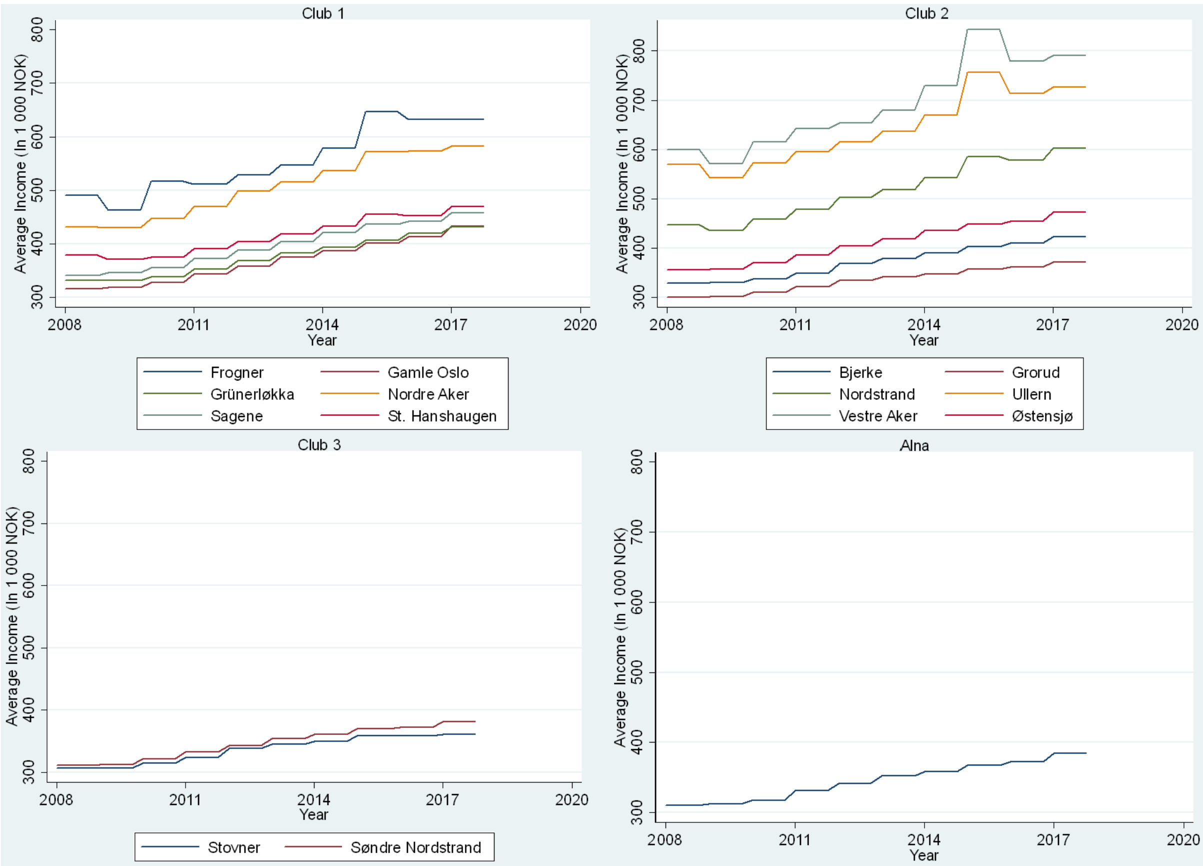

Figure A5.8: Level of income for the clubs identified by using property prices per square meter. Shows the average income of the population for each district at period t.

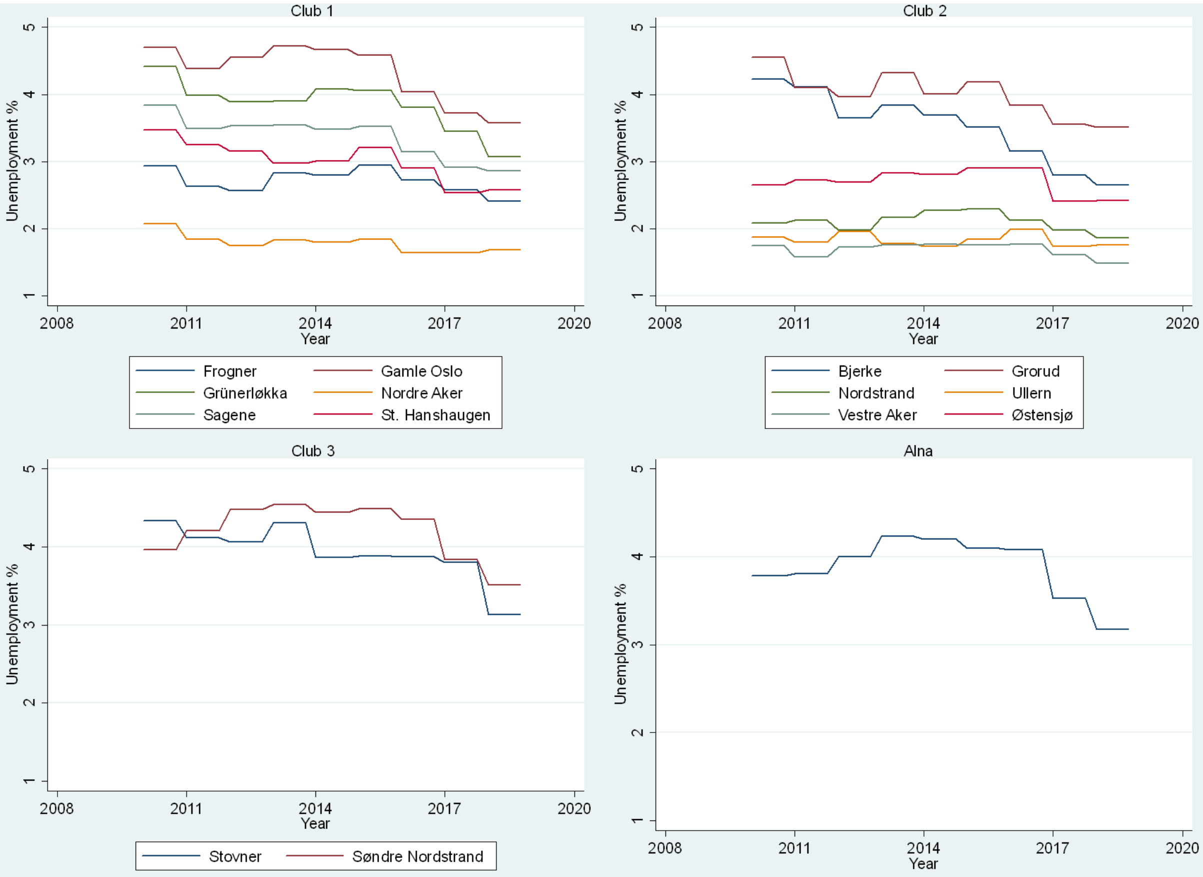

Figure A5.9: Unemployment rate for the clubs identified by using property prices per square meter. Shows the unemployment rate for each district at period t.

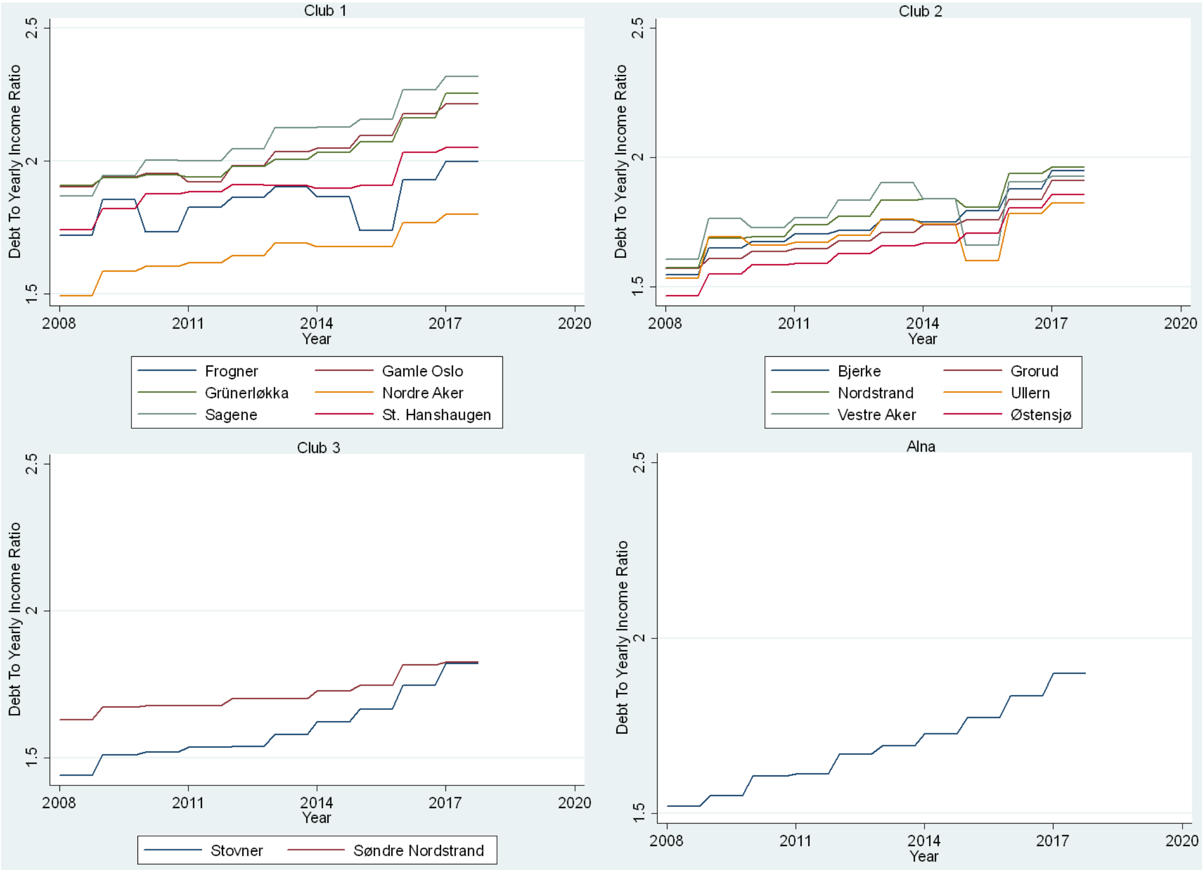

Figure A5.10: Debt to yearly income ratio in the respective convergence clubs identified by using property prices per square meter. Found by dividing the total average debt with the average yearly income. Shows the debt gearing of the population for each district at period t.

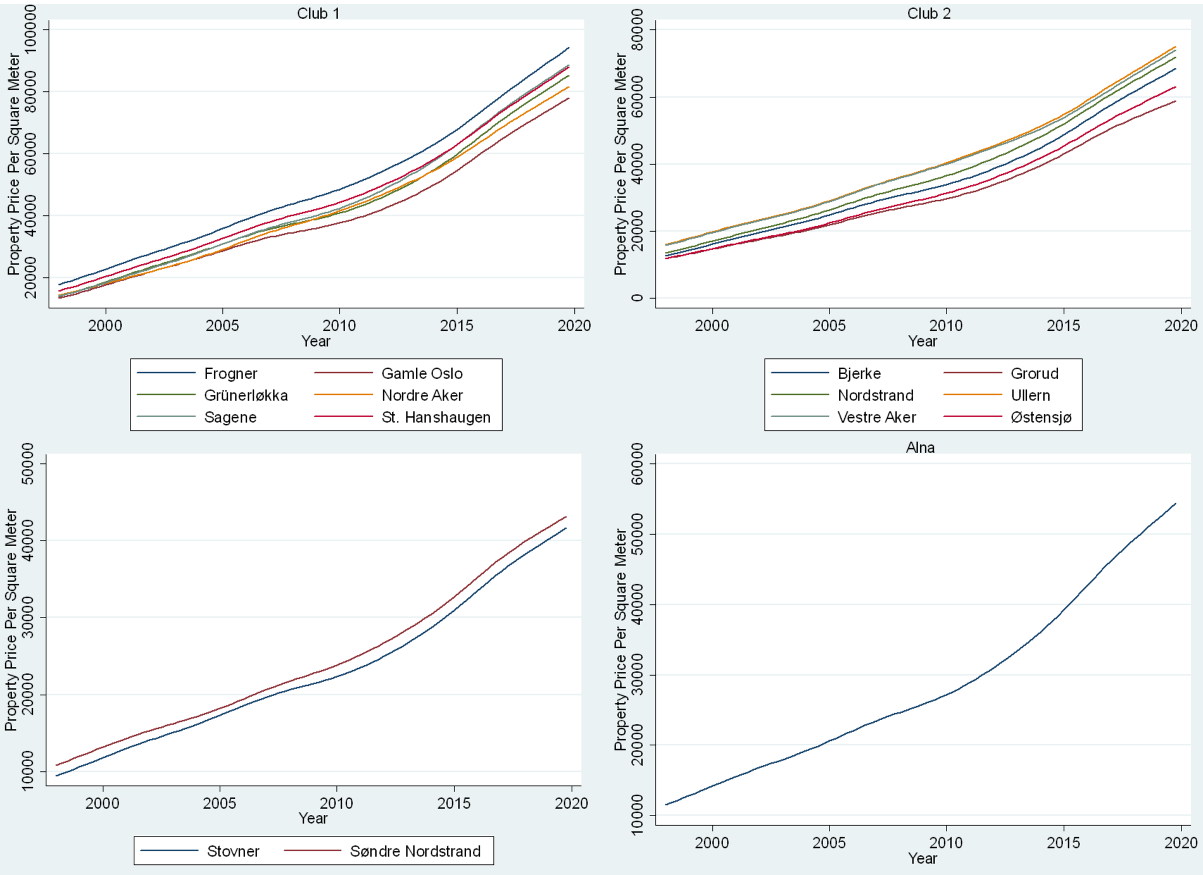

Figure A5.11: Graphs illustrating the price appreciation per square meter for all the clubs and their respective districts. Convergence clubs identified by using prices per square meter.

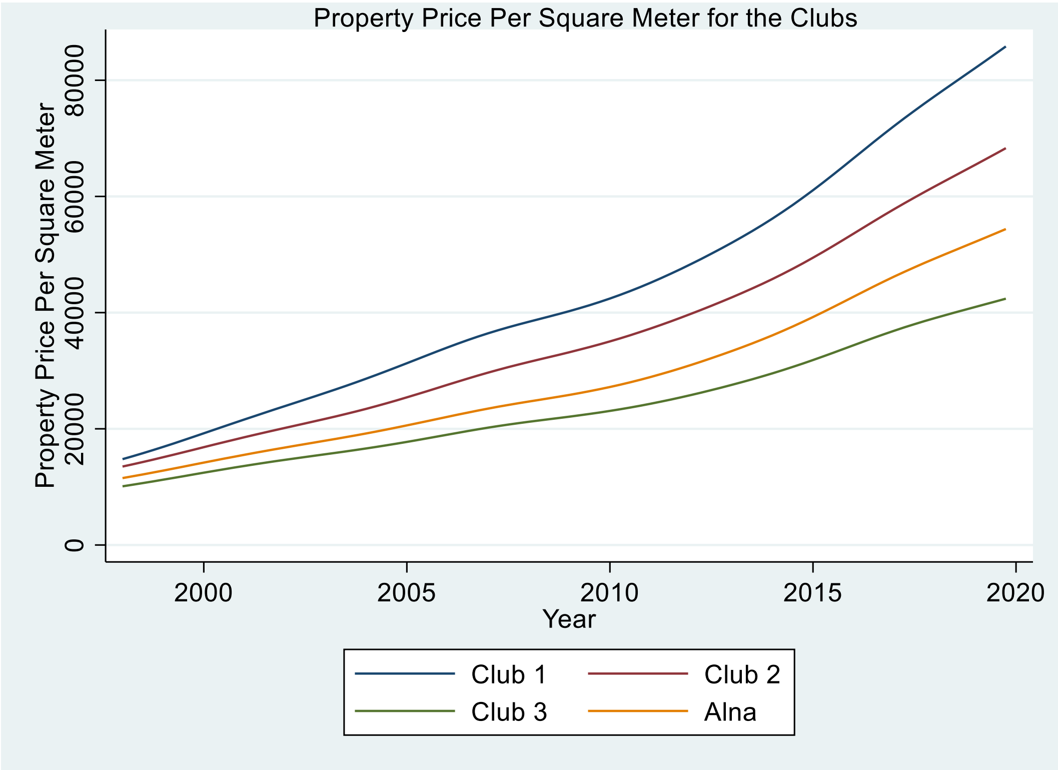

Figure A5.12: Graphs showing the average price growth per square meter for the convergence clubs in normal values. Convergence clubs identified by using prices per square meter.

Appendix - Tables

Residual²	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Quartals	.00044	.00002	23.55	0	.0004	.00047	***
Constant	.01419	.00044	32.15	0	.01332	.01505	***
Mean dependent var	0.02268	SD dependent var	0.08905				
R-squared	0.00452	Number of obs	122062.00000				
F-test	554.70926	Prob > F	0.00000				
Akaike crit. (AIC)	-244578.13924	Bayesian crit. (BIC)	-244558.71467				
*** *p*<.01, ** *p*<.05, * *p*<.1							

Table 3.1: OLS regression showing the relationship between the error term and the number of time periods between two sales.

Appendix - Theorem

In the regression property, $s_b^2 = \widehat{Ivar}(\hat{\varepsilon}_t)\left[\sum_{t=[rT]}^{T}\left(\log(t) - \frac{1}{T-[rT]+1}\sum_{t=[rT]}^{T}\log(t)\right)^2\right]^{-1}$.

$\widehat{Ivar}(\widehat{\varepsilon_t})$ is *heteroskedastic and autocorrelation consistent* (HAC), which means it allows for fitting of residuals that are both *heteroskedastic* and *autocorrelated*, and it is an estimation made by the residuals from Eq. 9:

(A1) $\quad \hat{\varepsilon}_t = log\left(\frac{H_1}{H_t}\right) - 2\log(\log(t)) - \hat{a} + \hat{b}\log(t)\,, for\ t = [rT], \dots, T$[27]

[27] Phillips & Sul, 2007, p. 1791.

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway