NHH

An empirical analysis of the Norwegian housing market

What drives the house price?

Walther Torset

Supervisor: Gernot Peter Doppelhofer

Master thesis in the financial program

NORWEGIAN SCHOOL OF ECONOMICS

Abstract

According to OECD (2018), Norwegian real house prices increased 44 percent from the end of 2008 until 2017, compared to an OECD average of 8 percent. Meanwhile, there has been a substantial change in several important aspects of the Norwegian households' financial base. In order to understand the house price development this thesis develops a house price model based on fundamental drivers of the housing market. Including debt in the model improves the in-sample fit, especially around the Financial Crisis. However, the model with debt is not robust. As a result the households' debt is not a part of the preferred model. In the out-of-sample period the real interest rate was close to zero, which made credit cheaper and increased expectations of future house price growth. As the preferred model emphasize the interest rate, it is able to accurately forecast the development of house prices in the out-of-sample period from 2015 to 2017.

In contrast to earlier studies, such as Jacobsen and Naug (2004) and Anundsen and Jansen (2013), the analysis finds that both the growth of house prices one year ago and one quarter ago impact the house price development. The importance of backdated growth implies that people observe the historic house price when they shape expectations to the future house price. This mechanism can cause a self-reinforcing house price spiral, which can lead to severe house price bubbles. Especially the effect of the growth in the preceding quarter is worrying, as it implies that people react faster to changes in house price growth, than what was found in the earlier studies of the Norwegian housing market.

The main drivers of the Norwegian house price are found to be the households' real income, the real interest rate, the housing supply, backdated growth in house prices and the households' expectations. According to the model, the real house prices in Norway were above their fundamental value from mid-2016 until the end of the sample in 2017.

Table of content

Table of figures

1. Introduction

House prices are an important factor in any modern economy, and a dwelling is by far the most valuable asset owned by a large share of households. The development of the house price is thereby closely tied to the households' wealth. When the house price change, it implies a wealth effect for the households, which in turn can impact their consumption (Jansen, 2009). An example of how the housing market is a central part in the modern economy and how it can have a severe negative impact, is the Financial Crisis in 2008 where subprime mortgages in the US market played a key role in building up the crisis (Barnes, 2009).

The spring and summer of 2017, Norwegian house prices started to decline after a period of dramatic growth (E24, 2018). In an interview with Business Insider from 2013, Jeff Gundlach and Nobel Prize winner Robert Shiller were asked about the Norwegian housing market. They characterized the market as "out of whack", and pointed out that Norwegian house prices had risen 77 percent from 2004 until 2013, compared to a 3 percent increase in the US. The Atlantic's Matthew O'Brien was also quoted in the article, saying that Norway's housing bubble "make ours look almost cute by comparison" (Perlberg, 2013). After 2013 came a period of low growth in house prices, before the prices once again started to rise. According to OECD (2018), the real house price in Norway have increased with 44 percent since the end of 2008, compared to an OECD average of 8 percent.

The households' decision of whether to buy a dwelling or not is likely to be impacted by many factors. Several aspects of the households' financial situation have changed substantially over the last years. Since 2014 Norwegian workers have had one of the periods with the lowest real wage growth since the second world war, partly due to a high growth in prices (Bjørnestad, 2017). The real interest rate on the other hand, was close to zero in 2016, which made credit much cheaper. In 12 months from August 2016, the debt of Norwegian households increased with 6.6 percent (Dagens Næringsliv, 2017). Due to the growth of the households' debt, the Norwegian government introduced several new mortgage regulations, which were intended to reduce the growth of debt. In order to better understand the essential relationship between the house price and its fundamental drivers this thesis develops an econometric model that can be used to analyze and understand the drivers of the Norwegian housing market. Can the Norwegian house price development be explained through fundamental drivers?

The paper is organized as followed: the first half of the paper presents relevant background information about the housing market and its potential drivers. Chapter 2 provides an overview of the house prices in Norway since the early 1980s. Chapter 3 takes a theoretical approach to what the fundamental drivers of house prices are. Chapter 4 examines the potential drivers' development the last three decades. Chapter 5 elaborates on important topics in time series analysis. The second half of the paper analyzes the relationship between the house prices and the drivers. Chapter 6 discusses the stationarity and order of integration of the relevant variables. In chapter 7 the house price model for the Norwegian economy presented by Dag Henning Jacobsen and Bjørn Naug in the article *"What drives house prices?"* from 2004 is re-estimated. However, the re-estimation on a sample almost double the size of the original finds clear evidence of serial correlation. In chapter 8 an alternative model is introduced. The model is built on the same variables as used by Jacobsen and Naug, but it allows for more dynamics to deal with the serial correlation. Variables considered in the theoretical discussion, that are not a part of the model already, are included to see whether it augments the model's performance. Chapter 9 presents the conclusion of the thesis.

2. The Norwegian housing market

After the Second World-War Norway was far from being a rich country by Western standards. The oil fields off the Norwegian coast was still to be discovered, and most households' survived on one income (Iversen K. O., 2016). During the 1960s, there was a desperate need of housing. The large baby boom generation that was born in the years following the war were maturing, and the supply of housing was insufficient. In addition, most young adults ended up living with their parents until they got married (Iversen K. O., 2016).

Figure 1: number of finished houses, annually (SSB, 2010)

Figure 1 shows how there was a rapid construction of new houses during the reconstruction of the country following the Second World War. Due to a limited housing supply the housing markets in the cities were regulated, and potential buyers had to register on waitlists for years before they were able to purchase their own home. In the 1960s the price of a house was usually around 3.5 times the annual income, with an interest rate three times higher than today (Iversen, K. O., 2016).

In the decades following the war, the Norwegian credit market had also been heavily regulated. From the late 1970s some modifications to the credit regulations were made, and when Kåre Willoch and his conservative party won the election in 1981, they started to speed up the deregulation (Regjeringen, 2013). The *additional reserve requirements* had constrained the supply of credit for decades. It was removed in January 1984, and the removal was intended to

be a permanent (Krogh, 2010). The regulations of the Norwegian housing market was abandoned September 1st 1982, and s a consequence, there was a quick rise in house prices the following months (Krogh, 2010).

In the early 1980s the Norwegian economy was booming. Following the deregulation of the housing market and the supply of credit, the Norwegian economy saw tremendous growth, mainly driven by credit financed consumption and investments. Norwegian banks were expanding dramatically, and opening plenty of new branches (Finansdepartementet, 1992). As can be seen from figure 2, the real house prices were building up until the first quarter of 1988. However, when the oil price fell in 1985-1986, the economy moved into a recession (Finansdepartementet, 1992). From 1988 to 1993 house prices dropped almost 30 percent, while the amount of annually finished houses dropped 47 percent. The period from 1987 until 1992 is called the Norwegian Banking Crisis. When Norwegians purchased a home in 1993, they paid on average two times their annual income. In addition, most households now relied on two incomes. Yet, the interest rate was moving up towards 15 percent (Iversen K. O., 2016).

Figure 2 compares the development of the real house price to the development of the real GDP since 1986. Since the Norwegian Banking Crisis in the early 1990s, the real house prices have been developing faster than the real GDP. From 1993 to 2016 the development in the real house price looks almost slightly exponential compared to the development in the real GDP.
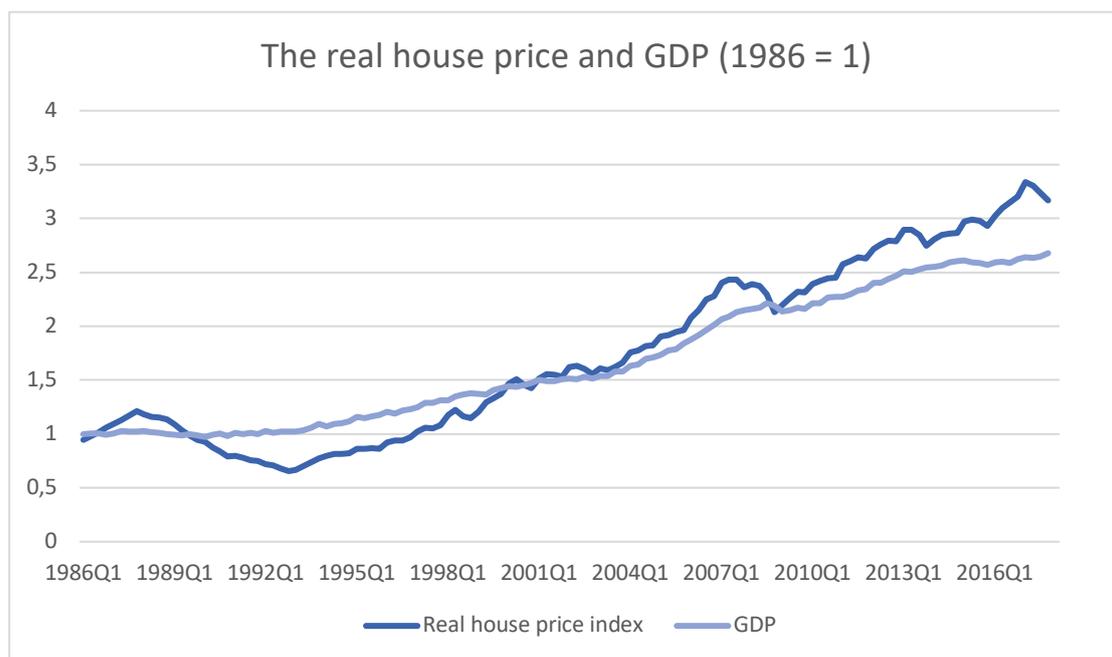
Figure 2: the real house price and the real GDP, deflated by CPI-ATE. (Eiendom Norge, 2018) and ssb.no

The four-quarter growth of the real house price was more or less constantly positive from the Banking Crisis until 2007, as can be seen in figure 3. Around the end of 2007, there was a lot of turmoil in international financial markets, due to large insecurity regarding American subprime loans. The collapse of the American investment bank Lehman Brothers on September 15th 2008 was the start of the international Financial Crisis (Barnes, 2009). The crisis had a dramatic, but short impact on Norwegian house prices. The real house price dropped 14 percent from 2007Q3 until 2008Q4. However, already in 2010Q3 the real house price had surpassed the prices from 2007. There were several reasons why the house price recovered swiftly. Among them were that the oil price and the world trade rebounded quickly, and the Norwegian central bank reduced the interest rate with 4,5 percentage points from October 2008 until June 2009 (Gustavsson, 2012).
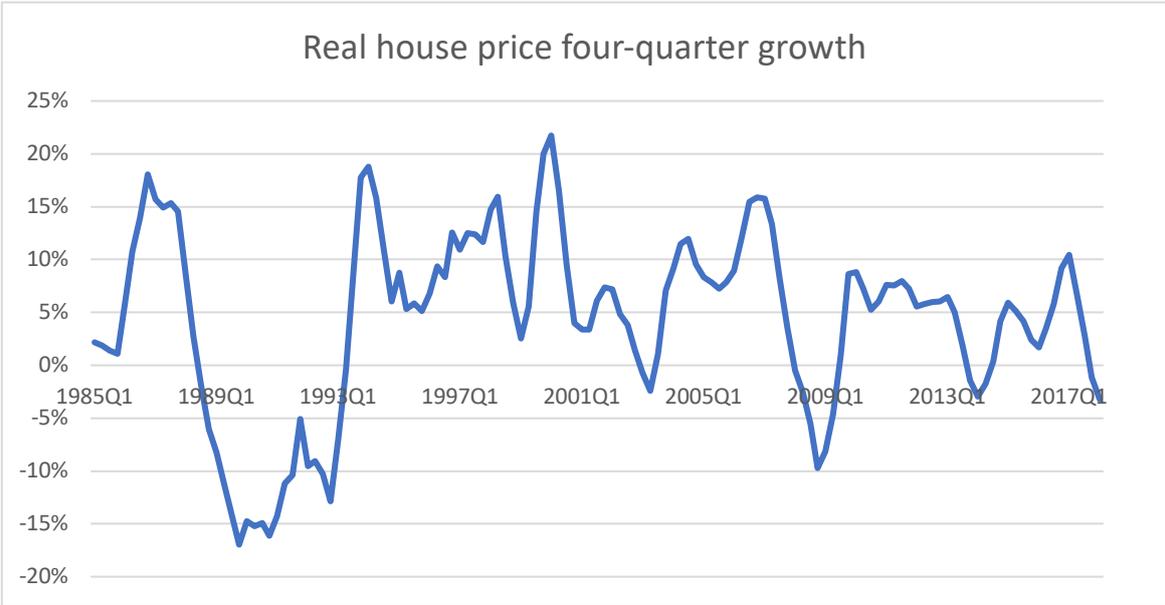
Figure 3: the four-quarter growth in real house prices since 1985

The average share of Norwegian households that owns their own dwelling over the last three years is 77 percent (SSB, 2018). Meanwhile, only 62 percent of the Swedes and 50 percent of the Danes owned their own dwelling in 2015 (Marschhäuser, 2015). The last years the average interest costs and installments on mortgages have in sum been higher on an annual basis than average rental costs (SSB, 2016). Why is there still such a high share of households that own their own dwelling? Tax advantages of owning a house in Norway is likely to be an important consideration, which is the topic of the next paragraph.

2.1 Tax advantages of housing

Not only has the interest rate been low since the Financial Crisis, but the Norwegian taxation system favors house purchase over house rental. Interest payments on mortgages is deductible from the tax bill, which is a big advantage over renting (Skatteetaten, 2018). In addition, if you lived in the dwelling during 12 out of the last 24 months, you do not have to pay tax of the income in case of a sale (Skatteetaten, 2018). In other words, paying on a mortgage is equivalent to saving in an investment object, where the return is free of tax costs. It is also possible to rent out up to half of your dwelling tax-free (Skatteetaten, 2018). In order for house renting to be beneficial, the house rental cost must be very low relatively to the cost of buying a dwelling due to the strong tax advantages (Iversen K. O., 2018) .

2.2 Regulations in the housing market since 2010

2010

The Financial Supervisory Authority (FSA) of Norway (Finanstilsynet) published a report in 2010, where they introduced new guidelines for a responsible supply of mortgages. The background for the new guidelines were according to the FSA that several Norwegian households were exposed to economic problems in case of a recession or a rise in the interest rate through their mortgages. As a measure to secure financial stability, they sought to create a more robust and sustainable practice in the housing and credit markets. The guidelines were intended to reduce the amount of large private loans, both relatively to income and the house price. The most important guideline was the recommended demand of a 10 percent equity share in the house purchase, so that only up to 90 percent of the total price of a dwelling could be debt-financed (Finanstilsynet, 2010).

2015

On July 1st 2015, new regulations came into action. It was expected that in order to cover the capital reserve requirements from Basel III in July 2013, the Norwegian banks would increase their margins. Yet, the house prices kept increasing. In 2014, 19 percent of mortgages had an equity share of less than 15 percent (Norges Bank, 2015). Following the ideas of macro-prudential policies, the Norwegian FSA sought to implement measures to make the Norwegian financial system more robust. The new regulations were a mandatory pre-testing of the households' ability to handle an increase in interest burden of 6 percent. More importantly, the banks' were not allowed to provide financing for more than 85 percent of the house cost.

However, the banks were allowed to deviate from the regulations in 10 percent of the loans in each quarter. The idea of this so-called "speed limit" for the banks were inspired by similar regulations in the UK and New Zealand (Norges Bank, 2015).

2017

In August 2016, house prices were 11 percent higher than income compared to the spring of 2008 (Finanstilsynet, 2018). Households with high debt could substantially decrease their consumption in case of a drop in house prices, which could work as an amplification mechanism in a potential recession. After the preceding regulations, house prices stabilized to a certain degree. However, the debt to income ratio kept increasing. The most important aspects were that potential buyers were only allowed to borrow up to five times their annual revenue. The growth in house prices had been especially high in Oslo in the previous years. As a result, the purchase of secondary dwellings in Oslo, could only be 60 percent debt-financed. This measure was aimed at an increasing speculation in the housing market. In addition, the "speed limit" was reduced to 8.5 percent for banks located in Oslo, while it was maintained at 10 percent for banks elsewhere (Finanstilsynet, 2018).

In 2018 the regulations were revised by the Norwegian FSA. Still, even with the new regulations on credit the growth in debt levels maintained high. The FSA assumed that the cap of debt-to-income ratio had contributed to a reduction in house price growth in 2017. Other important aspects that impacted the house price in 2017 according to the FAS was an increase in supply of housing, the population growth was diminishing and the real income growth had been, and was expected to be low (Finanstilsynet, 2018). Another important factor was the low interest rate in the previous years. It was not expected that the interest rate would drop further in 2017. From July 2018, the special requirements for Oslo will be taken away, and the speed limit will be reduced to 8.5 percent for banks all over the country (Finanstilsynet, 2018).

3. What are the fundamental drivers of house prices from a theoretical point of view?

The aggregated house price level is determined by housing demand and housing supply. According to Corder and Roberts (2008), the housing supply curve is steep and inelastic in the short-run due to adjustment costs as can be seen in figure 4. The production of housing is hard to scale as efficiently as a commodity due to factors such as availability of land, building permissions and the heterogeneity of the product. Therefore, the supply-side of the housing market is, in its nature, slow to respond to an increase in demand for houses. If there were no adjustment costs, there would be no difference between the long- and short-run housing supply curves, and the return from an additional house should equal the cost of building it (Corder & Roberts, 2008).
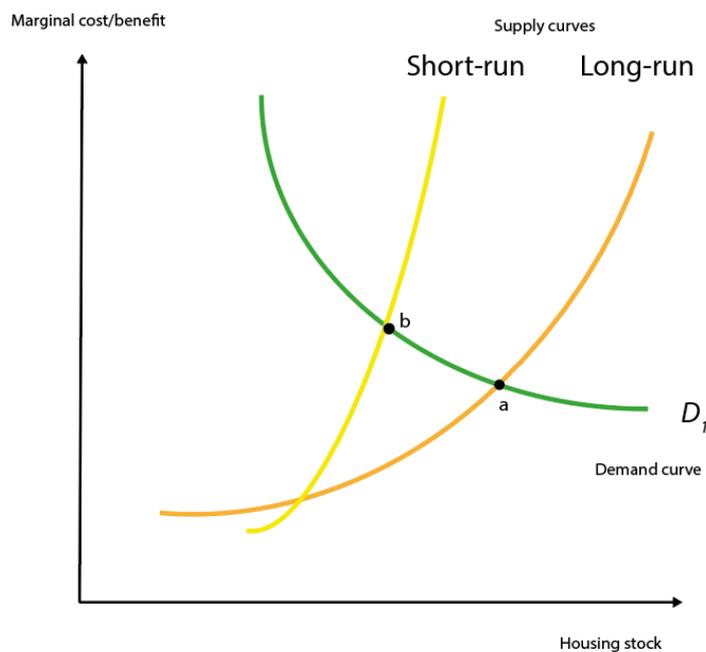
Figure 4: the short-run and the long-run supply curves of housing (Corder & Roberts, 2008)

In the short-run house prices are assumed to fluctuate with a change in demand for a given supply of housing. The housing supply is assumed to only impact the house prices in the long-run. For simplicity, this paper follows Hendry (1984), and assumes that changes in supply is given exogenously.

3.1 Housing demand

There are two components of housing demand, namely house purchases where the owner will live in the house, and housing as an investment object. 77 percent of Norwegian households owned their own dwelling in 2017 (SSB, 2018). Thus, a clear majority of the house purchases are for buyers to utilize the housing services themselves. Demand for owner-occupied dwellings will therefore be emphasized, and this demand is assumed to be proportional to the total housing demand. The economic framework presented below if based on the one proposed by Jacobsen and Naug (2004).

Housing demand can be explained through the following aggregate demand function:

$$H^D = f\left(\frac{V}{P}, \frac{V}{HL}, Y, X\right) \quad f_1 < 0, f_2 < 0, f_3 > 0 \quad (3.1)$$

Where H^D = housing demand, V = total housing costs for a typical owner – mostly the interest rate, P = index of prices of goods and services other than housing, HL = house rent, Y = households' real disposable income, and X = observable variables that capture effects from banks' lending policies, households' expectations, unemployment and demographic conditions.

From equation 3.1, we can see that the house price will decrease if housing costs connected with ownership goes up relatively to the house rents and the price of other goods and services. House prices will rise with increasing income, f_3. The households' disposable income is what constrains their purchasing power in a long horizon, as it is the basis on which they take up loans. Housing is a normal good, so the more people earn, the more likely they are to spend on housing. The variables in the vector X are important factors because housing is a consumer durable, the purchase of a dwelling is usually the most substantial purchase in the households lifetime and most households debt finance a large portion of the purchase.

Real housing costs can be defined as:

$$\frac{V}{P} \equiv \frac{PH}{P} BK = \frac{PH}{P}\left[i(1-\tau) - E\pi - (E\pi^{PH} - E\pi)\right] \quad (3.2)$$

Where $\frac{PH}{P}$ is the real house price, BK = housing cost per NOK invested in a dwelling, i = nominal interest rate, τ = marginal tax rate on capital income and expenses, $E\pi$ = expected inflation (expected rise in P measured as a rate), $E\pi^{PH}$ = expected rise in PH (measured as a rate).

The expression [i(1-τ)- Eπ] is the real-after tax interest rate. It represents the real interest cost of a housing loan, and the opportunity cost of investing in a dwelling. A higher interest rate will therefore increase housing costs. $(E\pi^{PH} - E\pi)$ represents the expected real rise in house prices. If the expected house prices are increasing, the cost of housing goes down, and owning a dwelling becomes more advantageous, which again leads to an increase in housing demand. If the house rents increase, it becomes more advantageous to invest in housing for rental purposes. Lower interest rates, and/or higher expectations to future house prices will also make it relatively better to invest in housing rather than leaving the money in a bank deposit. All these factors will drive up the demand for housing as an investment object.

Given a stable supply of housing in the short-run, the equation that defines the fundamental equilibrium house price (PH) is:

$$\ln PH = \beta_1 \ln HL_t + \beta_2 \ln Y_t + \beta_3 \ln BK_t + \beta_4 \ln H_t + \beta_5 g(X) + \varepsilon_t \quad (3.3)$$

where H = housing supply, and X = banks' lending policies (availability of credit), households' expectations, unemployment and demographic conditions.

House prices can fluctuate dramatically if the interest rate or other fundamental factors vary considerably. A higher unemployment rate leads to more pessimistic expectations to the future of the economy, as more people are becoming insecure about their future wage income and ability to repay debt. It also reduces the confidence in other's financial situation, which in turn can reduce the willingness to pay high house prices.

A reduction in interest rate on the other hand, will lead to expectations of a rapid rise in house prices. That can make potential buyers more eager to buy, and expedite their planned purchases, which will lead to a rise in prices initially, before the prices will start to fall again. A rise in house prices can be initiated through such a shock to one or more of the fundamental drivers. This will increase house prices, which again will make expectations of future price growth go up. Expectations of high future house prices makes the cost of housing lower and the advantages of owning higher. As a result, people are more inclined to buy, and this again leads to further price expectations. Eventually, such a process can drive house prices far beyond their

fundamental values. According to IMF, a positive and substantial deviation from the fundamental value of housing is an evidence of a housing bubble in the market (IMF, 2003). Yet, it is reasonable to assume that that the shock to the house price will disappear over time, and the house price will move towards its long-run fundamental value.

3.2 The banks' lending policies

Most households' debt finance a substantial share of the house purchase through debt. As a result, banks' lending policies can be important for the development in house prices. The lending policies depend on the banks' profitability, government regulations, customers' expected ability to repay debt and the collateral value of the households' dwellings.

$$L^s = h\left(O, REG, Y, U, \frac{PH}{P}\right) \qquad (3.6)$$
$$h_1 > 0, \quad h_2 < 0, \quad h_3 > 0, \quad h_4 < 0, \quad h_5 > 0,$$

where L^s = the banks' supply of credit to households, O = banks' profitability, REG = measure of government regulation of bank lending, Y = households' disposable income, U = unemployment rate, $\frac{PH}{P}$ = the real house price.[1]

From equation 3.6, we can see that banks will increase their supply of credit to households if the banks' profit go up, if the disposable income rise and if the real house price increase. On the other side, if the unemployment goes up or the government regulates the supply of credit more strictly, the banks will reduce their supply of credit. The households' debt will be used as a proxy to test for the effect of the banks' lending policies. Since the disposable income, unemployment rate and the real house price are already included directly in the model, the coefficient of the debt variable is only identifiable if the supply of credit is restrained by either the banks' profitability or by government regulations. The interest rate does not limit the banks' supply of credit directly. Instead an increase in the interest rate makes credit more expensive, and thereby it reduces the demand for credit.

---

[1] See Stiglitz (1992, Sections 6.2-6.3) for a theoretical discussion of the banks' lending policies.

Jacobsen and Naug found in 2004 that the supply of credit to households was constrained during the Banking Crisis in the early 1990s. On the other hand, they did not find that the supply of credit was limited by the banks' profitability after the crisis. They explain that it appears as if the supply of credit had a less independent effect after the crisis, than before and during the deregulation of the Norwegian credit market in the mid-1980s (Jacobsen & Naug, 2004b). According to Jacobsen and Naug, an insignificant effect of household debt on house prices in the house price model, indicates that lending was not limited by government regulations nor banks' profitability in their estimation period.

Other models for Norwegian house prices estimated on data from the 1980s and the 1990s found however a significant positive effect of household debt on house prices (Eitrheim, 1993) (Boug, Johansen, & Naug, 2002). Anundsen and Jansen found in 2011 a significant effect of debt in the multi-equation model for the Norwegian market (Anundsen & Jansen, 2013). A potential reason for this can be that they extended their data set until 1986, and thereby captured more of the effects from before the Banking Crisis, and in addition their series included the initial data from the Financial Crisis, where the supply of credit was likely reduced by the banks' profitability. In addition, IMF's house price equation for 18 OECD countries from 2004, and found a positive effect of credit growth on house prices from 1971 to 2003 (IMF, 2004). Similar results were found also in models for the Swedish and the UK markets (Hendry D. F., 1984) (Barot & Yang, 2002). Credit was regulated in large parts of the estimation periods of these models.

4. Empirical examination of fundamental drivers

In the following chapter, the development in the fundamental drivers from chapter 3 are examined. An overview of the data can be found in appendix A. The deregulation of the Norwegian housing market in the mid 1980s lead to structural changes in the market. Therefore, the analysis will start from 1986, post-deregulation. The paper utilizes real values in the analysis. Real values provide a better image of the actual development in the housing market over time, as the variables are corrected for inflation. All monetary variables have been deflated with the consumer price index adjusted for tax changes and excluding energy products (CPI-ATE), also called the core inflation. House prices are unlikely to be appreciably affected by short-term fluctuations that are due to tax changes or energy prices.

4.1 The households' real income

The data used to measure the households' disposable income in the paper is the households' wage income, as tax-motivated equity income has had a considerable effect on the measured developments in the households' disposable income. The fluctuations in share dividends probably had a small impact on the household demand for dwellings.
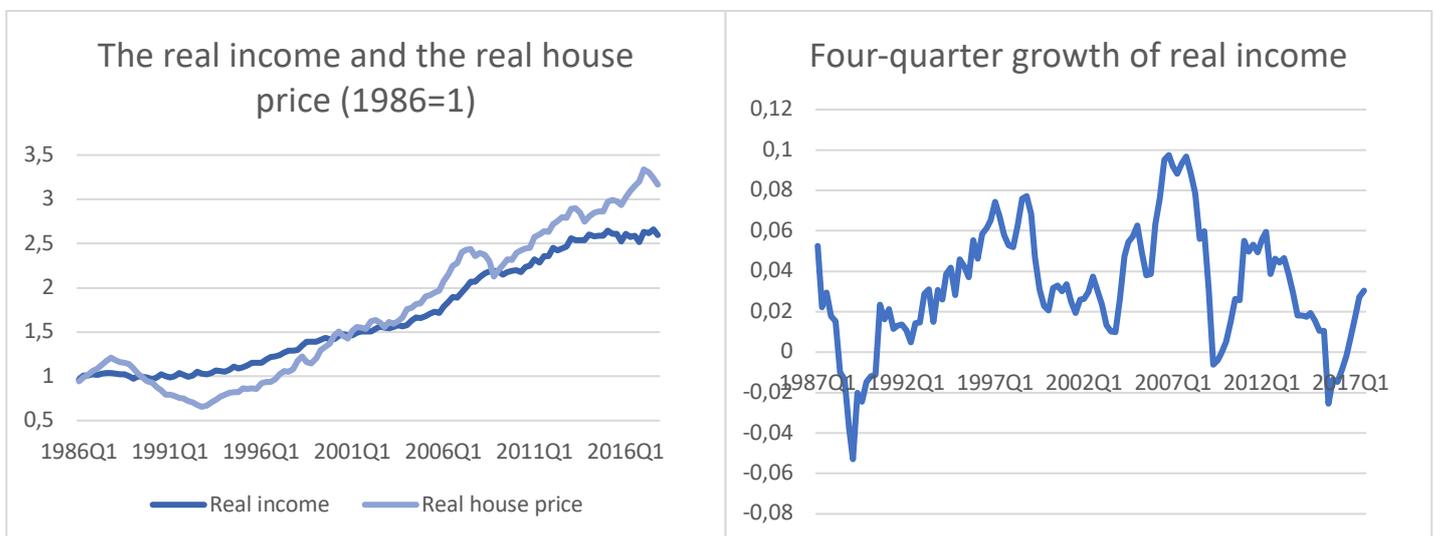
Figure 5: the real income and the real house price.

wage growth was negative momentarily from 1988 to 1990 during the Banking Crisis, before there was an increasing growth until 1999. In the period from 2000 until 2003 the world economy was in a period of recession, following the events such as the Dotcom bubble, and the invasion of Afghanistan and Iraq (World Bank, 2017). The state of the world economy impacted

the Norwegian economy, and the growth of real income was low. The real income increased dramatically from the beginning of 2004 towards the Financial Crisis. In 2008, however, the income dived from extremely high to negative growth. Since 2009 the house prices have outpaced the real income, which can be a dangerous situation, and not sustainable in the long-run. The increase in the house price must be financed, and when the income is insufficient, the households must take up more debt.

In Norway, the organized share of the workforce was reduced from 50.5 percent in 2006 to 49 percent in 2016 (Birkelund, 2017). Once a year the labor union (LO) and the Confederation of Norwegian Enterprises (NHO) negotiate the employees conditions (NHO, 2018). The labor union rarely accepts a reduction in the nominal income, unless the companies are in a very bad situation. A change in the income will therefore be permanent in the Norwegian economy. However, the price development have the last years contributed to a low income growth, even though the nominal wage has been rising (SSB, 2018).

4.2 Real interest rate after tax and the households' real debt

The tax deductibility has been subtracted from the real interest rate to account for the tax advantages of owning a house. The interest rate is the price of credit, and a low interest rate will make mortgages cheaper. In addition, households observe the interest rate and use it to shape expectations to the future. The data material of debt in this paper is represented by the households' total debt to Norwegian banks.
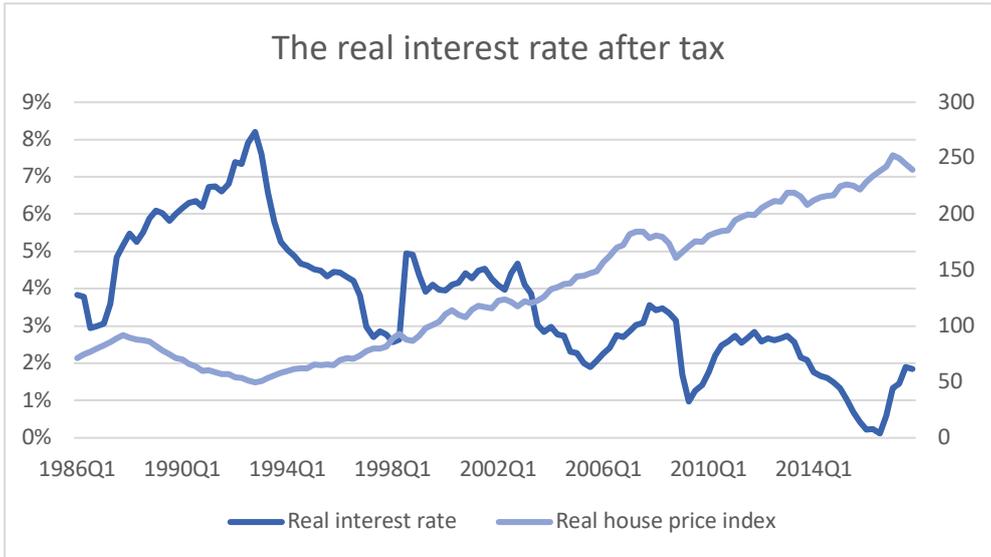
Figure 6: the development of the real interest rate after tax (deductibility)

During the late 1980s and the 1990s, the real interest rate was in general much higher than in the succeeding decades. During the same period, there were two extensive tax reforms. Among different changes, the interest deductibility in the taxes was cut from 45 percent to 40 percent in 1978. Then, in 1992 it was further reduced from 37 percent to 28 percent (Christensen, 2018). The reforms were intended to make it less beneficial to take up debt. Since the tax deductibility was reduced, it contributed to higher values for the real interest rate after accounting for the tax.
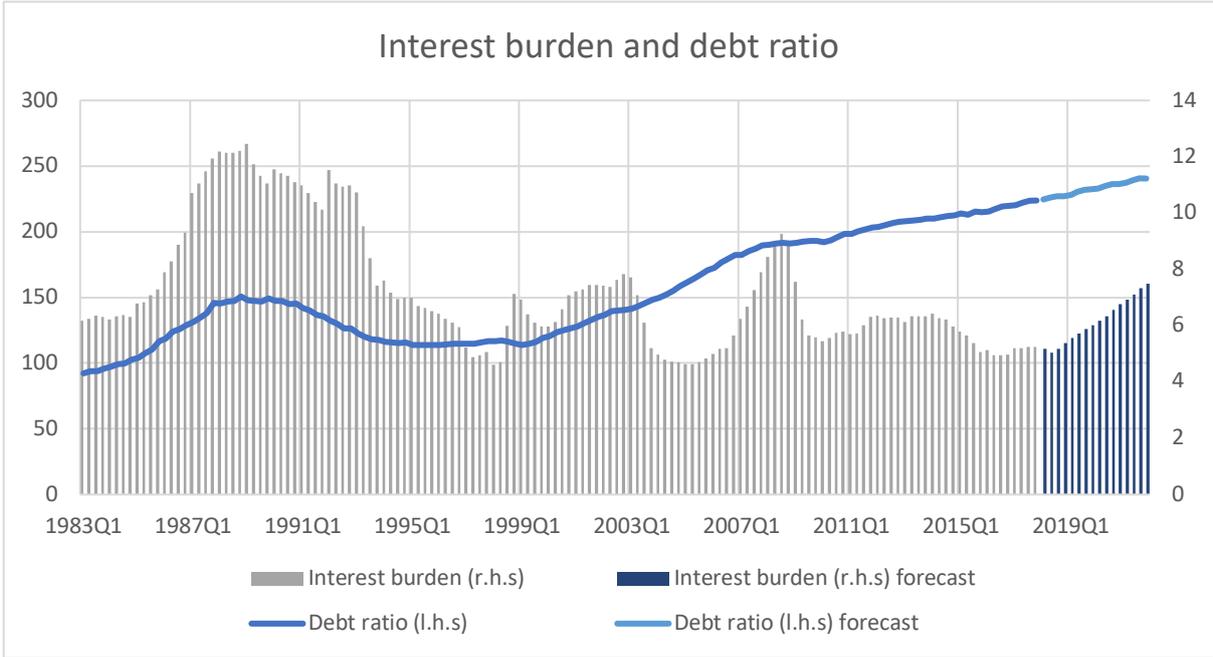
Figure 7: the households' nominal interest burden and nominal debt ratio. (Norges Bank, 2018)

Figure 7 provides an overview of the households' interest burden, which is calculated as interest expenses as a percentage of disposable income and interest expenses, and the households' debt ratio, which is loan debt as a percentage of disposable income. The interest burden was clearly high during the whole period of the Banking Crisis and the Financial Crisis, when the central bank was trying to cool down the economy. However, when the crisis eventually hit the Norwegian economy, the Central Bank dramatically reduced the interest rate. Since the central bank kept the interest rate low in the years following the crisis the interest burden has remained low. Meanwhile, the households' debt has been increasing relatively to the households' income. From the discussion of the real income, we saw that the real house price had outpaced the real income since 2009. The increasing debt ratio is another indication of that households' debt finance their purchases to a larger extent. The regulations on mortgages does not seem to have had a strong effect yet. However, figure 8 shows that following the new regulations from

January 2017, the real debt stabilized for a couple of months, before it started to rise again. It will be interesting to see how the regulations will impact the debt levels in a longer horizon. From the forecasted values, we can see that the Central Bank expects the interest rate to rise. Given the high levels of debt compared to income, the interest burden will be substantial for many households.
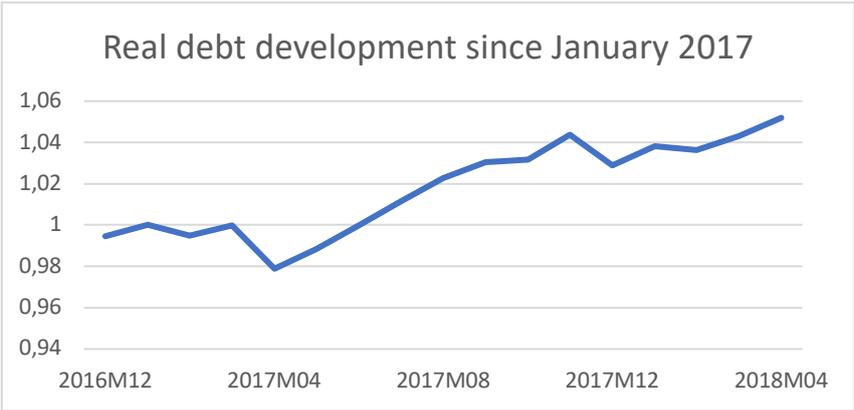
Figure 8: the real debt's development since January 2017. (2017M1 = 1)

The households debt was growing to high levels before the Banking Crisis. When the crisis hit, the growth in debt fell sharply. Following the Banking Crisis came a period with exponential growth in the debt, leading up to the Financial Crisis. The one-year growth in the second quarter of 2006 was over 16 percent. From the indices of the real debt and the real house price, we can see that the household debt grew much faster leading up to the Financial Crisis, and it fell dramatically after the crisis. The growth in household debt was strongly negative from 2008 until 2013. Since 2013, the growth in debt has again been positive.
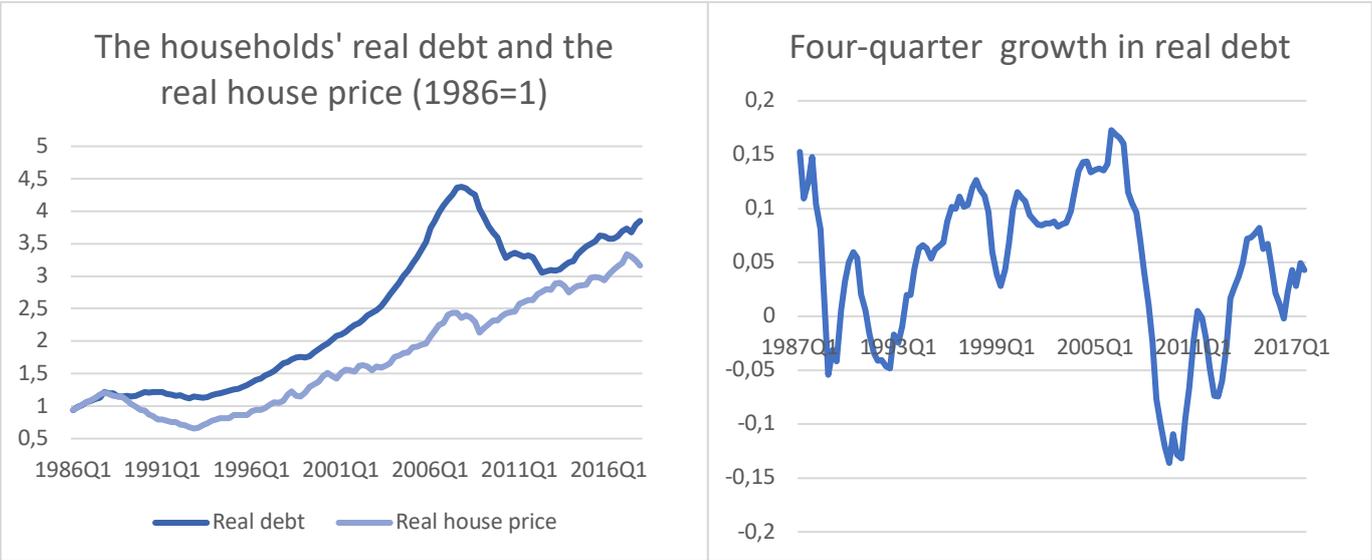
Figure 9: the development in the real debt, and the four-quarter growth in the real debt

4.4 Unemployment rate

The unemployment rate is an important indicator of the activity in the economy.[2] The unemployment rate was rising dramatically during and following the Banking Crisis in the early 1990s. As mentioned before, the world economy was moving into a recession after 2000. As the unemployment is closely related to the development of the GDP, the unemployment was high in the same period. Then the unemployment was declining to low levels before the Financial Crisis, before it jumped from around 2 percent to 3 percent following the crisis. From 2008 the unemployment has been stable around 3 percent.
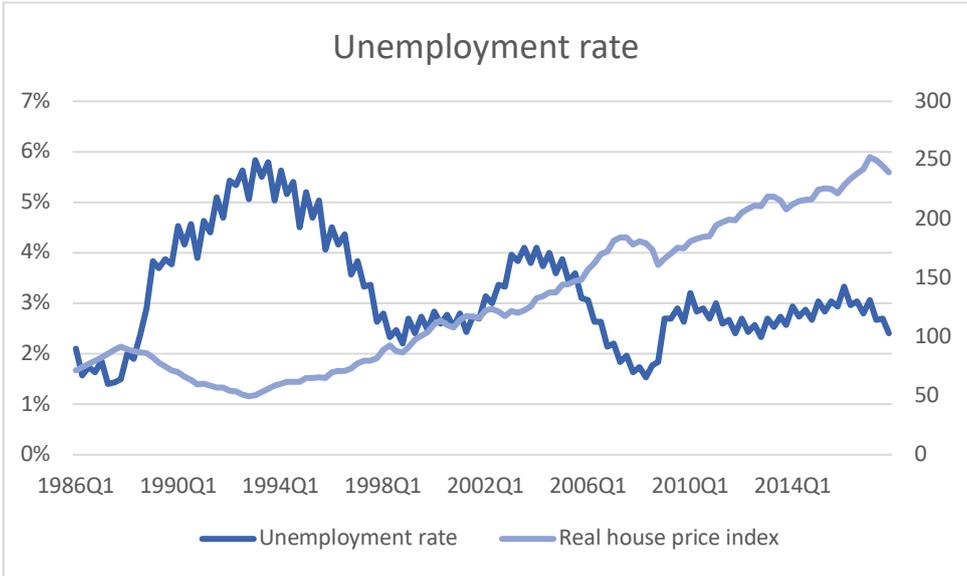
Figure 10: the unemployment rate

4.5 Population effects

An issue with the analysis of demographic effects on the house price is that most available data exist only in an annual frequency. The quarterly growth in the population is however available from the last quarter of 1997.

---

[2] Hence the concept of Okun's law, which is an empirically observed relationship between the unemployment and the GDP. For more information see *Okun, Arthur M. (1962). "Potential GNP, its measurement and significance". Cowles Foundation, Yale University.*
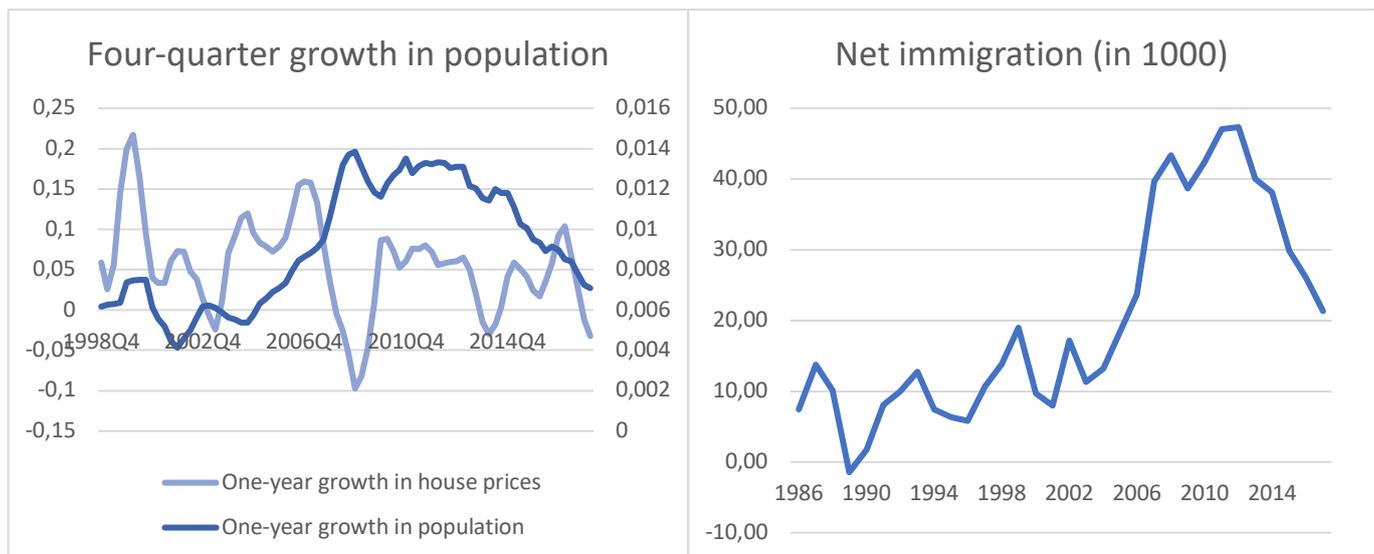
Figure 11: four-quarter growth in population and the net immigration

The population was increasing from 2000 until the Financial Crisis, which was mostly driven by a strong positive net immigration. Immigrating adults have an immediate need for a dwelling, different from a newborn baby that lives with its parents. Thus, an increase in the population through immigration can put a higher pressure on the housing market in the short-run, than an increase in population due to a higher birth rate. According to the Norwegian FSA, the diminishing growth in population was one of the potential reasons for why the house price declined in 2017 (Finanstilsynet, 2018). It is however odd that it would impact the markets so suddenly, considering that the growth has been falling sharply since 2014.

A slow moving demographic effect is that the average amount of people in a household has been going down the last 50 years. In 1960, there were 3.3 people on average in a household. In 2017, the number was reduced to 2.2 people per household.[3] For a given population this effect will increase the amount of households, and increase the demand for dwellings. In addition, the share of the population that lives in urban areas has been rising steadily, from 70 percent in 1990 to 81 percent in 2017. The urbanization is likely to have an impact on the house prices in the cities, as it results in a higher demand in limited areas.

---

[3] The numbers for the demographic changes have been collected from SSB.no, and the sources can be found in appendix A.

Another demographic aspect is that the share of inhabitants in the "establishment phase" has been going down. The "establishment phase" is the period in the life when people get their first job, settle down and buy a dwelling. In 1990 the share of people between 19 and 34 years was 25 percent. The share declined to 20 percent in 2008 before it increased slightly again to 22 percent in 2017. The average age of the first-time buyer was 28 years of age in 2017 (Pihl, 2017). As the share of the population in the establishment segment has been going down, that implies that there are less first-time buyers, and is an argument against price appreciation.

4.6 House rents

SSB measures the development of house rental costs in Norway by registering the real house rent on established and new rental contracts (Johannessen, 2004). The house rent has been growing steadily over the last decades. During the Banking Crisis the four-quarter growth in the house rents dropped from almost 9 to 0 percent. Since 2000 the growth in the house rents have been varying around 3 percent. Thus the house rents seems to be much more stable then several of the other potential drivers. The house rental costs have been developing much more steadily the last decades than the real house price, according to the graphs in figure 12.
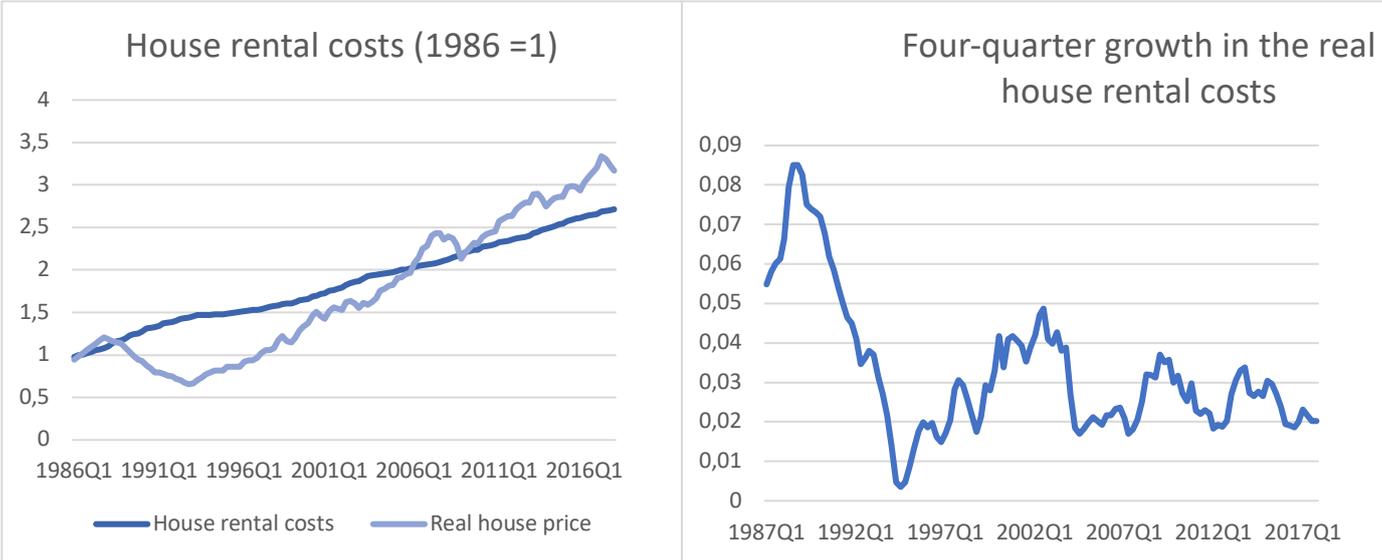
Figure 12: the real house rental costs

4.7 The households' expectations

The households' expectations is an indicator created by TNS Gallup and Finans Norge since 1992. It measures the households' expectations to the state of the country's and their own

economy in the future. It is created through a survey consisting of five questions, and then the answers are accumulated into a number between -100 and 100. A number below zero represents a degree of pessimism, while a number above zero represents a degree of optimism. In this paper the variable has been divided by 100, so that the value is between -1 and 1. The purpose of the variable is to capture psychological effects in the market.

The households' expectations are fluctuating over time, as can be seen in figure 13. Yet, there seems to be a positive mean value. The expectations are only strongly negative during the Banking Crisis and the Financial crisis, and in addition the expectations were negative during 2016. The period followed the fall in the oil price in 2014 and 2015, and the growth of income was going down. Highly educated people were especially worried about the future of their own economy. However, it seems like factors such as a low interest rate quickly turned the expectations around, and in 2017 the expectations were positive again.
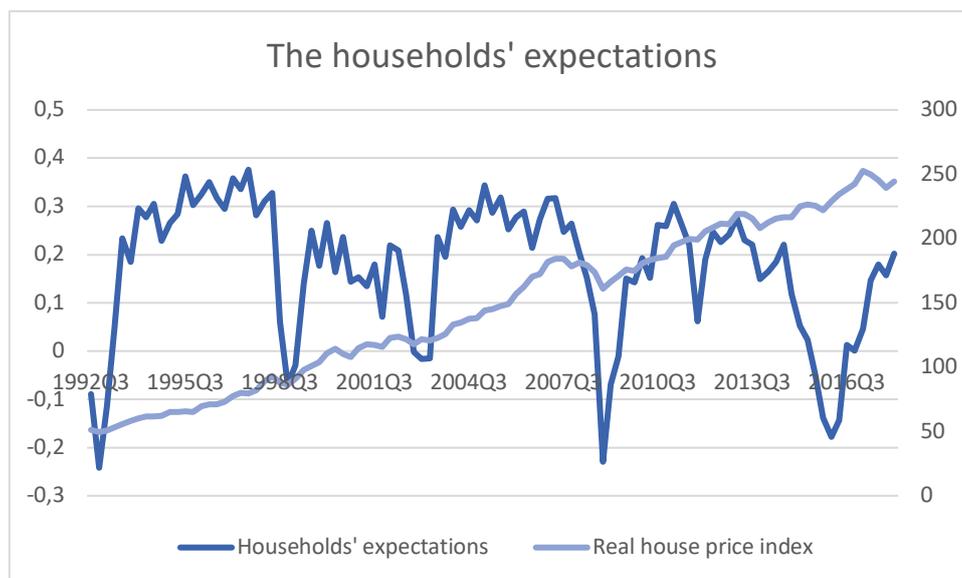
Figure 13: the households' expectations and the real house price

4.8 Supply of housing

The housing stock is calculated through a perpetual inventory method, where the capital depreciation is assumed to be 2 percent over time. The housing stock measured in fixed prices fell sharply during the Banking Crisis. However, it was almost back at the same growth rate as before the Banking Crisis in 2007. When the Financial Crisis hit, it once again dropped sharply. From 2010 the housing stock has had an increasing growth rate.

25
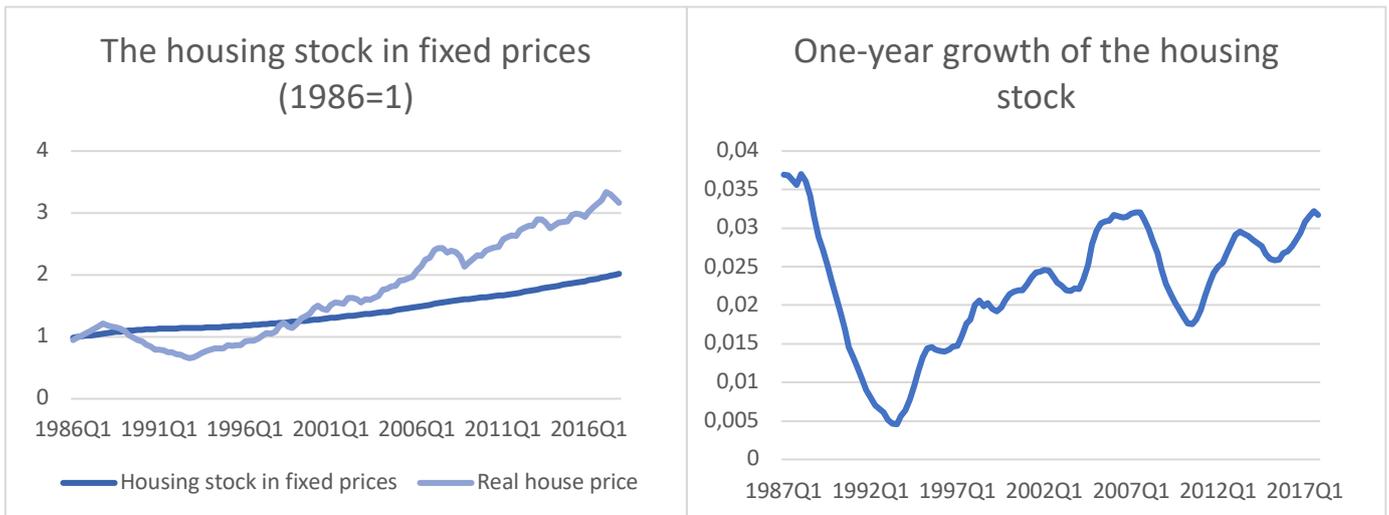
Figure 14: the housing stock in fixed prices

Another measure of the housing supply is the initiated square meters of housing.
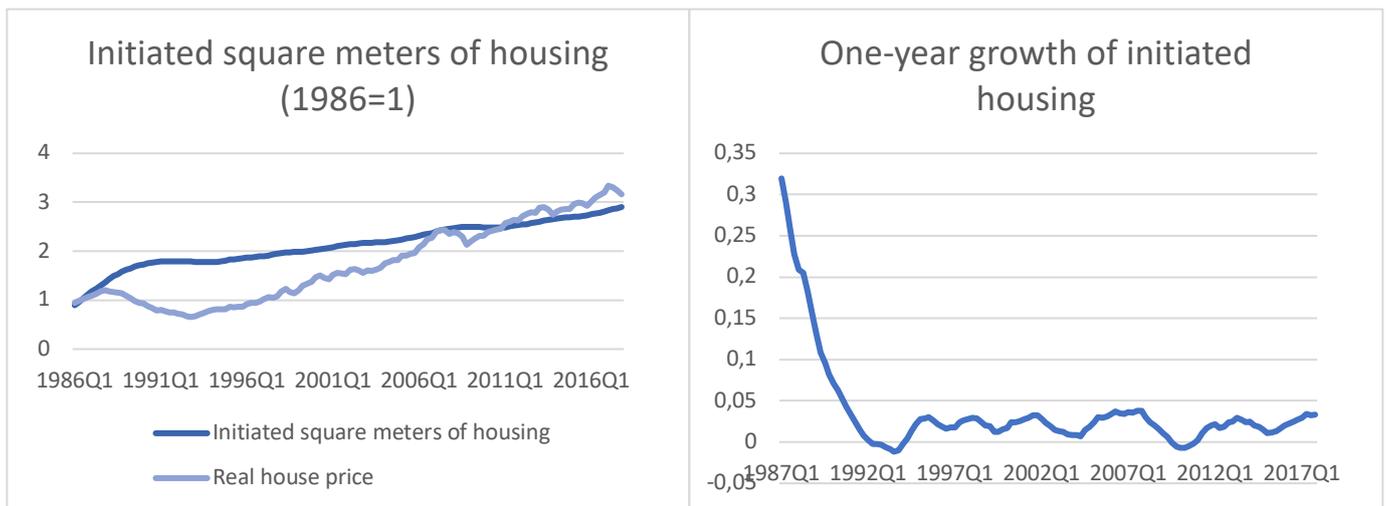
Figure 15: the initiated square meters of housing.

There was a construction boom after the deregulation of the housing market in the early 1980s up towards the Banking Crisis. After the crisis hit the economy, there was a dramatic drop in the growth of construction similar to the drop in the housing stock. The growth in initiation of new square meters of housing fell from 30 percent in 1987 to -1 percent in 1994. Since the Banking Crisis the growth of the initiated square meters of housing has been positive, except for a short period in 2010 following the Financial Crisis.

5. Important concepts in time-series analysis

5.1 Serial-correlation

Time series is a set of repeated observations over time of the same variable. A new and important consideration to make when we want to work with time series, rather than cross-sectional data, is the presence of serial-correlation. Serial-correlation is when a variable at period *t* depends on the value from its last period, *t-1* (Bjørnland & Thorsrud, 2015). To determine the degree of serial-correlation, we can graphically examine the data, which can give us an idea of whether the data is serial-correlated or not. Traditionally the Durbin-Watson has been a popular way to check for serial-correlation. However, the Durbin-Watson test can be indecisive. Therefore this paper emphasizes the Breusch-Godfrey test and the Ljung-Box test for serial-correlation[4].

5.2 Stationarity

An important type of time series is the autoregressive (AR (*p*)) process, where *p* represents the amount of lags of the dependent variable we include. AR (*p*) relates the value of variable y_t to y_{t-1} through the correlation coefficient ρ and a random disturbance u_t that is white noise i.i.d ~ N($0, \sigma_u^2$) (Bjørnland & Thorsrud, 2015).

$$y_t = \rho y_{t-1} + u_t \quad (5.1)$$

An AR (0) process will only depend on the error term u_t, while an AR (1) process depends on the last period's value and the error term. If $|\rho| < 1$ then an AR(1) is covariance stationary with finite variance, even though it depends on the last periods value, as any shock from the past will gradually disappear over time. In order for any least squares estimation on time series to be valid, we need to be sure that the stochastic properties of our data is stationary. Stationary data is characterized by a constant mean $E(y_t) = \mu$, constant variance $Var(y_t) = \sigma^2$ and a covariance that depends on the distance *s*, and not the time *t*, $Cov(y_t, y_{t-s}) = \gamma_s$ (Woolridge, 2013). Any time series that fail to satisfy these criterias, are what we call non-stationary. Several economic variables tend to be non-stationary in levels (Bjørnland & Thorsrud, 2015).

---

[4] For technical details of the tests, please see *Breusch, T. S. (1978). "Testing for Autocorrelation in Dynamic Linear Models". Australian Economic Papers*, and *G. M. Ljung; G. E. P. Box (1978). "On a Measure of a Lack of Fit in Time Series Models"*

A *random walk* is a time series process that only depends on its present value and white noise errors (Bjørnland and Thorsrud, 2015):

$$y_t = y_{t-1} + u_t \quad (5.2)$$

For a random walk, an initial shock will never die out, so any future value depends on the historical development of the series. If the residual is left alone on one side, the expression becomes stationary, as u_t is i.i.d $\sim N(0, \sigma_u^2)$. Neither its mean nor variance depends on time (Bjørnland and Thorsrud, 2015). A variable that becomes stationary after differencing it once y_t, is called integrated of first order I(1).

$$y_t = \alpha + y_{t-1} + u_t \quad (5.3)$$

If we add a constant term to the random walk formula, we get a random walk with drift. A random walk with drift's deviations from the deterministic trend are not stationary. This implies that a shock will be permanent, and the deviation will be permanent (Bjørnland and Thorsrud, 2015). Using non-stationary time series, such as a random walk, in normal regression analysis might lead to spurious results. We can evaluate the stationarity of our data through graphical analysis, a correlogram, a Dickey-Fuller test or a Phillips-Perron test.

The Dickey-fuller test takes the difference of the particular time-series once and get:

$$\Delta y_t = \alpha_0 + \alpha_1 t + \theta y_{t-1} + u_t \text{ where } \theta = (\rho - 1) \quad (5.4)$$
$$H_0: \theta = 0 \text{ and } H_A: \theta < 0$$

If $\rho = 1$, the series is a random walk. If $\theta < 0$, then $\rho < 1$ and we have a stationary time-series (Bjørnland & Thorsrud, 2015). The trend t can be included to account for a potential deterministic trend. If a trend is included, then $H_0: \theta = 0$, which implies that y_t is a random walk with drift versus $H_A: \theta < 0$ where y_t is trend stationary (Bjørnland and Thorsrud, 2015).

With the **augmented Dickey-Fuller** (ADF) **test** we can test for serial-correlated residuals by allowing more lags of the dependent variable. The testing procedure is the same as the ordinary Dickey-Fuller test, but it is applied to the model:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \cdots + \delta_{p-1} \Delta y_{t-p+1} + \epsilon_t \quad (5.5)$$

where α represents a constant, t represents a time trend and p is the lag order of the autoregressive process (Bjørnland and Thorsrud, 2015). The ideal lag order can be found through information criterias.

To deal with a higher order of serial-correlation than what is assumed in the ordinary Dickey-Fuller test, the Phillips-Perron (PP) test makes non-parametric correction to the t-statistic. The test is thus robust against unspecified serial-correlation and heteroscedasticity in the disturbance process of the test model (Phillips & Perron, 1988). The PP test performs worse than the ADF test in finite samples, as it is based on asymptotic theory, and thus works better in large samples (Davidson & MacKinnon, 2003). Both tests have disadvantages such as poor small sample power, which can lead to unit root conclusions even if that is not the case. If we suspect that our data has a structural break we can use a Zivot-Andrews test, that allows us to test for a structural break and a unit root simultaneously[5].

5.3 Cointegration

By making non-stationary data stationary through differencing, we potentially lose valuable information in the data, since the causality between different time series often are described in the level version of the data. However, using non-stationary data in a regression can lead to spurious results. Therefore, it is essential to confirm that there exists a cointegration between the data in levels. The presence of cointegration between the variables in levels can help specify econometric models that aligns with economic theory (Bjørnland and Thorsrud, 2015). Even if two time series are non-stationary, the difference between them can still be stationary. From equation 5.6 we can see that, if the error term u_t is stationary, then by definition the difference between the two variables are stationary too, since u_t is i.i.d $\sim \mathrm{N}(0, \sigma_u^2)$ (Bjørnland & Thorsrud, 2015).

$$u_t = y_{1,t} - \beta_1 y_{2,t} \quad (5.6)$$

The statistical properties of the linear combination are the same as for a I(0) variable, and we say that the two series are cointegrated (Bjørnland and Thorsrud, 2015).

---

[5] For more technical details please see *Andrews, D., Zivot, E. 1992. Further evidence on the Great Crash, the oil price shock, and the unit-root hypothesis. Journal of Business and Economic Statistics 10, 251-70*

The Engle-Granger cointegration test takes the assumption that if two series are integrated of order I(1), they might cointegrate, and the unknown cointegration coefficient has to be inferred from the data (Engle & Granger, 1987). To test whether the residual is stationary, we can use an ADF test. The critical values for the test will depend on whether or not we include a constant or trend in the cointegration regression (Bjørnland and Thorsrud, 2015).

Johansen's test allows for the potential existence of several cointegrations simultaneously. We can calculate the appropriate rank, or cointegrating relationships, with the help of a trace test statistic. The trace statistic works in the way that it first tests a null hypothesis that the appropriate rank is zero, and the alternative hypothesis that there are more than zero cointegrated relations. If the test cannot reject the null hypothesis, it moves over to test a null hypothesis where the appropriate rank is one. It continues until it finds a trace statistic that is smaller than the critical 5% value, which will yield the correct rank.[6]

5.4 Error-correction models (ECM)

Cointegration represents the long-run equilibrium relationship between the variables. It describes how two variables drift together. The variables might have temporary deviations from the long-term values, but we can allow fluctuations around the equilibrium by letting the equation have some dynamics. Error-correction models allow both long-run and short-run dynamics enter the same model, through the use of cointegration.

5.4.1 The Engle-Granger two-step method

Equation 5.7 represents an error-correction model, and is only internally consistent if the two variables are cointegrated.

$$\Delta y_{1,t} = \beta_1 \Delta y_{2,t} - (1 - \rho)(y_{1,t-1} - \beta_1 y_{2,t-1}) + u_t \qquad (5.7)$$

---

[6] The purpose of the Johansen's test in this paper is to confirm whether there exists one or more cointegrations between the long-run variables. The Johansen's test will not be explained in detail. For more information on the Johansen's test, the reader is adviced to see «Modelling of cointegration in the vector autoregressive model» by Søren Johansen, Economic Modelling (2000)

The error correction term is defined as $EC_{t-1} = (y_{1,t-1} - \beta_1 y_{2,t-1})$, and $(1 - \rho)$ represents the adjustment term. As both the left side and the first term on the right side of the equation above are stationary, the model will be inconsistent if the EC_{t-1} is not stationary[7]. Thus, after we have estimated the error-correction term, we have to check whether the residual is stationary with an ADF test. The long-run relationship is maintained through a mechanism that corrects any deviations from the long-run equilibrium. If $y_{1,t-1} > \beta_1 y_{2,t-1}$, then the EC_{t-1} term is positive. Yet, as $\rho < 1$ by assumption, the total effects will be negative. Thus, if one variable such as house prices is above its long-run equilibrium level relative to income, the error correction mechanism will move house price growth downwards, until the equilibrium level is again restored (Bjørnland and Thorsrud, 2015).

For richer dynamics over time, it is possible to use a more general version of equation 5.7 where we include more lags:

$$a(L)\Delta y_{1,t} = b(L)\Delta y_{2,t} - \gamma(y_{1,t-1} - \beta_1 y_{2,t-1}) + u_t \qquad (5.8)$$

where we have the lag polynomials of a(L) and b(L), allowing for a more dynamic structure determined by the data (Bjørnland and Thorsrud, 2015).

Engle-Granger's two-step method has empirical weaknesses. First of all, the unit root test and cointegration test lacks power. In addition, we can get biased estimates in small samples, and we cannot do any inference on the cointegrating variables in the first step, as the equation contains non-stationary variables (Brooks, 2005). Yet, the two-step model is popular as it is intuitive, easy to understand and the model is super consistent[8].

5.4.2 The one-step method

In the one step method we do not have to estimate the EC term before we run the error-correction regression. Instead, we calculate all coefficients simultaneously. The long-run adjustment speed will be represented by the coefficient of the lagged dependent variable. The significance of the lagged dependent variable is used to tell whether we have a cointegration or

---

[7] If $y_{1,t}$ and $y_{2,t}$ are I(1) variables, then $\Delta y_{1,t}$ and $\Delta y_{2,t}$ are I(0) variables, and stationary.

[8] Super consistency is when the OLS-estimate for two I(1) variables will move towards the true value much faster than OLS with stationary variables only, when the residual is stationary and the sample grows. This comes from the I(1) variables dominating the I(0) variables asymptotically.

not. Not only does the one-step model have higher power than the two-step model, but it also has much less biased estimates in small samples, as we model the short-run dynamics and the long-run relationship at the same time. Inference about the coefficients in the long-run is valid, given that they are cointegrated, which is a big advantage over the two-step method (Davidson & MacKinnon, 1993).

5.4.3 The autoregressive distributed lag model and Bounds testing method

This section follows Giles (2013) that has made an intuitive overview of the ARDL and Bounds testing methodology developed by Pesaran, Shin, & Smith (1999). The ARDL model can be defined as:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + \alpha_0 x_t + \alpha_1 x_{t-1} + \cdots + \alpha_q x_{t-q} + \varepsilon_t \quad (5.9)$$

where y represents the dependent variable, x represents the explanatory variable(s) and ε_t is a random disturbance term with no serial correlation. In order for ARDL to take the error-correction form, the ARDL equation is reformulated into:

$$\Delta y_t = \beta_0 + \Sigma \beta_i \Delta y_{t-i} + \Sigma \gamma_j \Delta x_{1,t-j} + \Sigma \delta_k \Delta x_{2,t-k} + \theta_0 y_{t-1} + \theta_1 x_{t-1} + \theta_2 x_{t-1} + \varepsilon_t \quad (5.10)$$

which is called an unrestricted ECM. The difference from a two-step ECM is that the error-correction term in step two has been replaced by $\theta_0 y_{t-1} + \theta_1 x_{t-1} + \theta_2 x_{t-1}$. In fact, until this stage, the estimation is highly similar to the one-step ECM.

According to Pesaran, Shin and Smith (1999), the cointegration literature sometimes put too restrictive assumptions on the cointegrating regressors x_t to be integrated of order one or more. The motivation behind Pesaran et al.'s ARDL framework was to prove that there could be a long-run relationship between two or more variables y_t and x_t, even if all the regressors x_t under consideration were I(0) (Pesaran, Shin, & Smith, 1999). The reason is that under the null hypothesis of the Bounds test that there exists no long-run relationship between y_t and x_t, the y_t process will be I(1) irrespective of whether the regressors are I(0), I(1) or mutually cointegrated. Pesaran et al. (1999) highlights that the ARDL and Bounds testing method is particularly relevant when there is doubt concerning the order of integration of variables such as the unemployment rate and the interest rate in a long-run relationship with other variables that are clearly I(1).

In order to estimate an ARDL model, none of the variables can be integrated of second order I(2), as it would invalidate the methodology. The amount of lags for the differenced variables are chosen automatically by the software with AIC. Further, there should most importantly be no serial correlation, but also no heteroscedasticity and normally distributed residuals. If the residuals are well behaved, the Bounds test can be used to check for a cointegrating relationship. The Bounds test is simply a F-test on the long-run coefficients from equation 5.10 with the hypothesis:

$$H_0\colon \theta_0 = \theta_1 = \theta_2 = 0, \text{against } H_A\colon \theta_0 = \theta_1 = \theta_2 \neq 0 \quad (5.11)$$

A rejection of H_0 implies that we have a long-run relationship between the variables. However, exact critical values does not exist for a test with a mix of variables that are potentially I(0) or I(1). Pesaran, Shin, & Smith (1999) provides bounds on the critical values for the asymptotic distribution of the F-statistic. They supply upper and lower bounds of the critical values. The lower bound is based on the assumption that all variables are I(0), and the upper bound is based on the assumption that all variables in question are I(1). There is an indecisive area between the lower and upper bounds. If the test statistic is neither above the upper bound or below the lower bound, but in between, the test is inconclusive.

As a potential cross-check we can perform a Bounds t-test, where $H_0\colon \theta_0 = 0$ against $H_A\colon \theta_0 < 0$. If the t-statistic for the lagged dependent variable is greater than the upper t-test bound provided by Pesaran et al. (2001), it is another evidence for a long-run relationship between the variables.

If the critical value exceeds the upper bound, the Bounds test concludes with cointegration. Then we can estimate a meaningful *long-run equilibrium relationship* between the variables:

$$y_t = a_0 + a_1 x_{1t} + a_2 x_{2t} + v_t \quad (5.12)$$

As well as the usual error-correction form:

$$\Delta y_t = \beta_0 + \Sigma \beta_i \Delta y_{t-i} + \Sigma \gamma_j \Delta x_{1,t-j} + \Sigma \delta_k \Delta x_{2,t-k} + \varphi z_{t-1} + e_t \quad (5.13)$$

$$\text{where } z_{t-1} = y_{t-1} - a_0 - a_1 x_{1t} - a_2 x_{2t}$$

A potential weakness of the single equation error-correction models is if there are in fact more than one long-run relationship in the data. Then a multi-equation approach would be appropriate.[9]

---

[9] However, Pesaran et al. (2001), footnote 30, page 26, adds "clearly the system approach developed by Johansen (1991) can be applied to a set of variables containing a possible mixture of I(0) and I(1) regressors. But in such cases the result of the trace or the max eigenvalue test will be difficult to interpret, as it will not be possible to identify whether the reduced rank outcome (if any) is indicative of the existence of long run relationships or is due to the presence of I(0) regressors in the model." During the initial work of the paper, the data was examined with the Johansen's test. However, the test did not find more than one cointegration between the variables in question. Therefore, I chose to proceed with the ARDL framework, as it provided a more intuitive and parsimonious model.

6. Unit root analysis

The following chapter examines if the variables to be used in the further analysis have unit roots. Statistical unit root tests can have low statistical power in finite samples (Cochrane, 1990). Therefore the conclusion of the variables' stationarity will be based on the expectations to the variable, the graphical analysis and the unit root tests. In addition, as long as none of the variables are integrated of second order I(2), it does not matter for the Bounds test whether the *true* order of the integration of a variable is I(0) or I(1) when testing for a cointegrating relationship.

In the rest of the paper, lowercase letters indicates that the variable has been log transformed. All variables have been log-transformed, except the real interest rate and the households' expectations. A summary of the unit root test results can be found in Appendix B.[10]

6.1 Real house price

Financial theory would argue that the fundamental price of a house could be related to the stream of future (rental) income, much like how the price of stocks reflects discounted future dividends. If this stream of income is affected by permanent shocks this would carry over to house prices in levels. Thus, I expect the real house price to have a unit root. The ADF test and the PP test cannot reject non-stationarity, and the differenced real house price is clearly stationary according to all tests.[11] I conclude with the real house prices being I(1).

---

[10] For the statistical analysis I have used the software packages Stata 15 SE and EViews 10 Student Version. Stata is the standard statistical software used at NHH, and it is a multi-purpose statistical software, with its strengths particularly in cross-sectional data and panel data analysis. This became evident while working with advanced time-series models. EViews was better suited for time-series analysis, but it was time consuming to get familiar with a new software.

[11] Including a structural break for the dramatic fall during the Banking Crisis does not change the conclusion of the real house price being I(1).
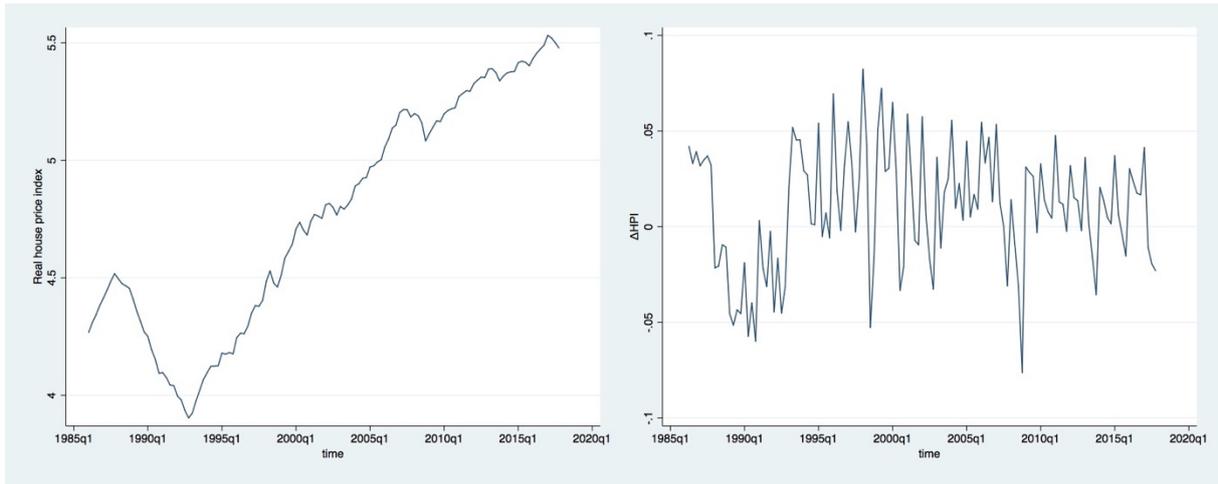
Figure 16: the real house price and the differenced real house price

6.2 Real income

While the growth in real income has slowed down the last years, the growth has been more or less positive constantly since the Banking Crisis. As the economy grows, there are more values to be distributed as wages. The accumulated real wage in the economy is unlikely to return to a mean value over time. For the last decades have not only the Norwegian GDP seen a tremendous growth, but the population has increased too. Thus, a shock to the households total wage income is expected to be permanent, and the variable is non-stationary in levels. The expectations are in accordance with the unit root tests, which confirms that the real income is I(1).
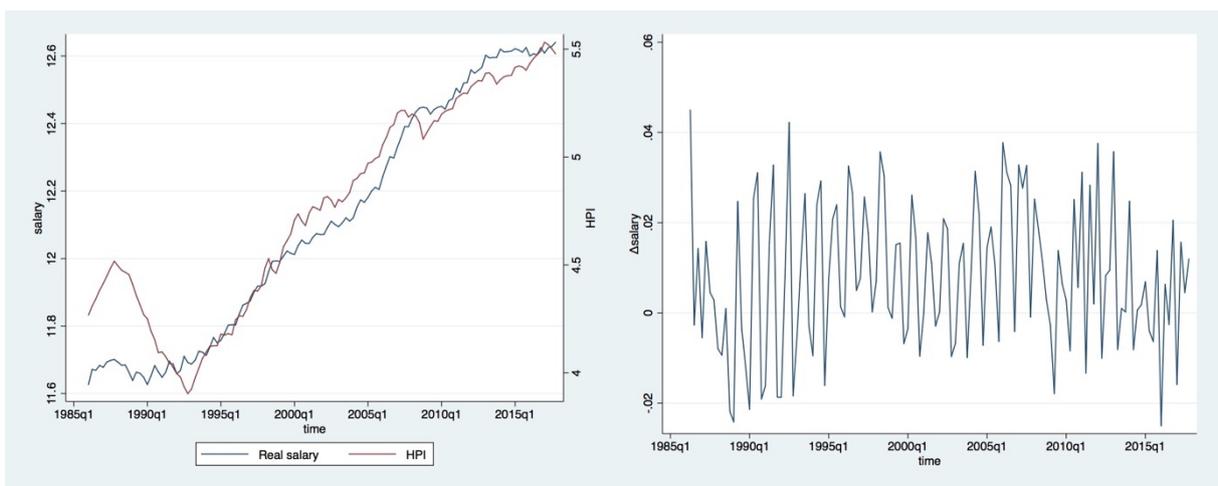
Figure 17: the households' real income and the differenced income

6.3 Real interest rate after tax

It is unlikely that the interest rate will rise or decline indefinitely. The central bank will rise the interest rate when the economy is heating up, and it will lower the interest rate when the economy is slowing down. If the interest rate is too high over a longer period, it could dramatically decrease the activity in the economy. Therefore, a significant drop or rise in the interest rate is usually temporary. However, if the interest rate is in fact stationary, it would be a highly persistent series, where shocks die slowly out. It can take years before the Central Bank choose to change the interest dramatically. For instance, the examination of the interest rate in chapter 4 found that it had been continuously low since 2008. It looks like the interest rate in figure 18 is trending downwards, and it resembles a random walk with drift. Since the interest rate seems to be clearly persistent in this sample I expect the data to be non-stationary in levels and stationary in differenced form. The differenced interest rate in figure 18 looks clearly stationary. We can see how the big deviations are quickly returning to the mean. The unit root tests confirms the expectations, and I conclude with the interest being I(1).
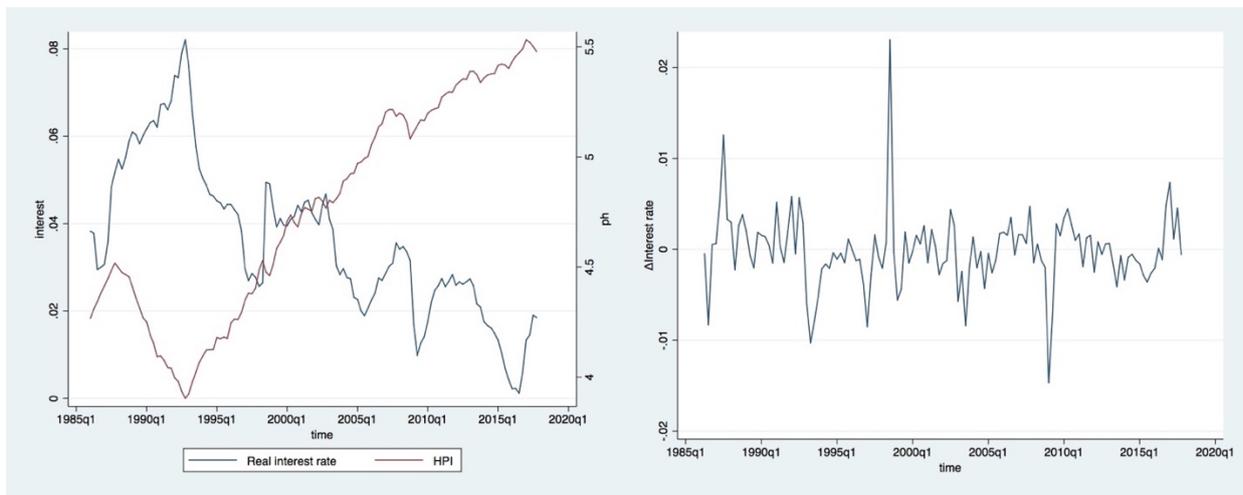
Figure 18: the real after tax interest rate and the differenced interest rate

6.4 Real debt

Households can easily take up debt, as long as they satisfy all of the bank's requirements. Paying down on debt, however, usually takes a lot of time. Most households do not have large cash reserves available. The only way they could repay most of their debt in the short-run would be if they sold their dwelling. Thus, it is expected that changes to the households' debt is highly persistent and non-stationary. The unit root tests confirms that the data is non-stationary in

levels.[12] The ADF and the PP tests concludes with the differenced data being stationary. Thus, the households' real debt is integrated of order one I(1).
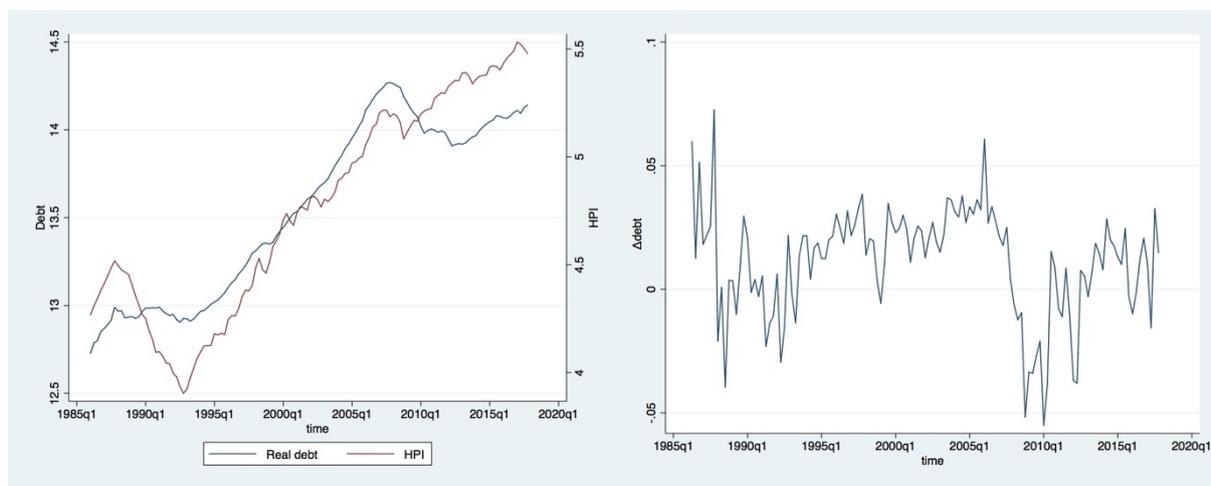
Figure 19: the households' real debt and differenced real debt

6.5 Unemployment

It is not unlikely that the unemployment is stationary in the long-run. If the unemployment rate rised indefinitely, the economy would more or less stop. If a majority of employees lost their income, total consumption would go down. That would force other companies to lay off more people, and so on. Through this mechanism it could amplify the effects in of a recession. Therefore, governments and central banks pay close attention to the unemployment rate. If it reaches alarmingly high values, the government will usually introduce measures in order to make the unemployment rate to go down again. On the other hand, there are certain frictions in the labor market, which makes it very hard to have zero unemployment. As a consequence, the basic unemployment rate is a non-zero value, also called the natural rate of unemployment. Over time, the unemployment rate should vary around a long-term mean value. From figure 20, the unemployment rate in levels seems to have weak stationary characteristics. The differenced unemployment rate looks clearly stationary. The ADF test concludes with stationarity for the unemployment rate, while the PP test concludes with non-stationarity. Both tests conclude that the differenced unemployment rate is stationary. In sum the conclusion is that the unemployment rate being I(0).

---

[12] Including a structural break for the financial crisis does not change the conclusion of non-stationarity.
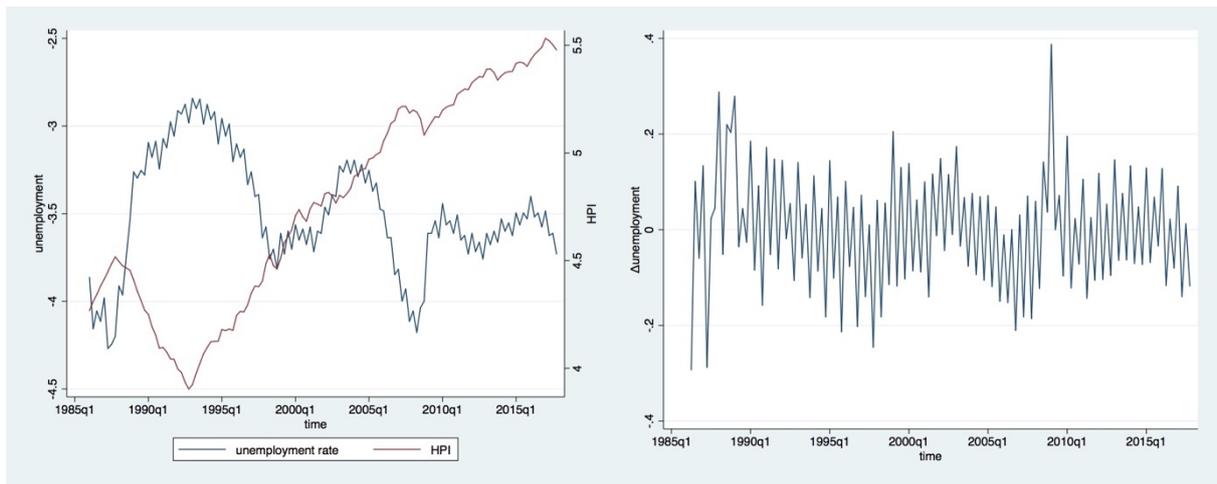
Figure 20: the unemployment rate and the differenced unemployment rate.

6.6 House rents

If the demand for house rentals go up the landlord will be able to increase the rents. When the rental cost has been increased there is no mechanism driving the rents back to a mean value. In addition, rental contracts are by law allowed to adjust for the general price growth once a year. Therefore shocks to the house rents are most likely highly persistent. The unit root tests confirms that the house rents are non-stationary in levels, and difference-stationary. Thus, the house rent is integrated of order one I(1).
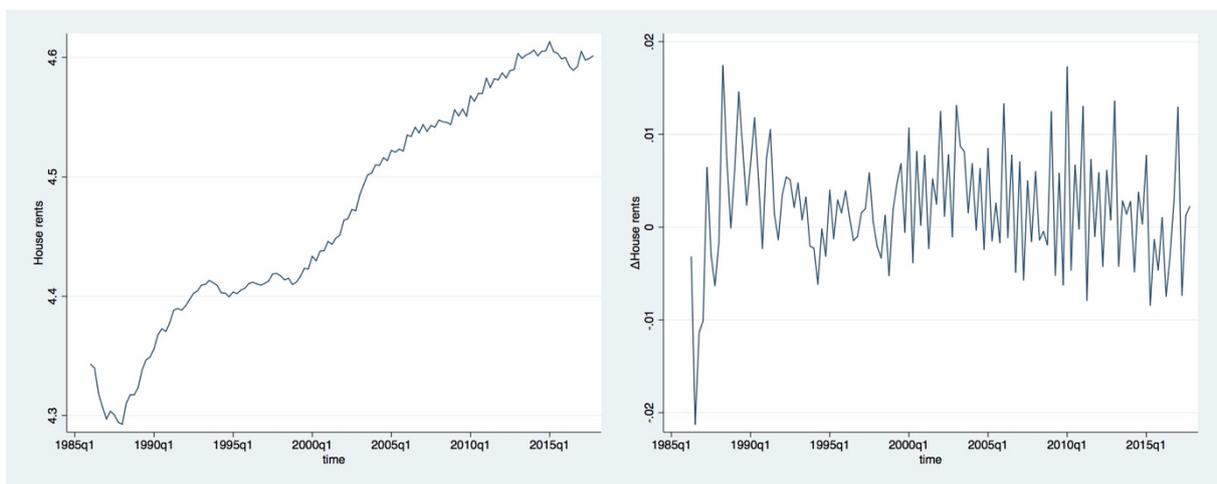
Figure 21: house rents and differenced house rents

6.7 The households' expectations

Psychological expectations go quickly up and down, often following big events in arenas such as the political landscape or in financial markets. Therefore I expect the variable to be

stationary. Both graphs in figure looks stationary, which is confirmed by the unit root tests. The conclusion is therefore that the households' expectations are I(0).
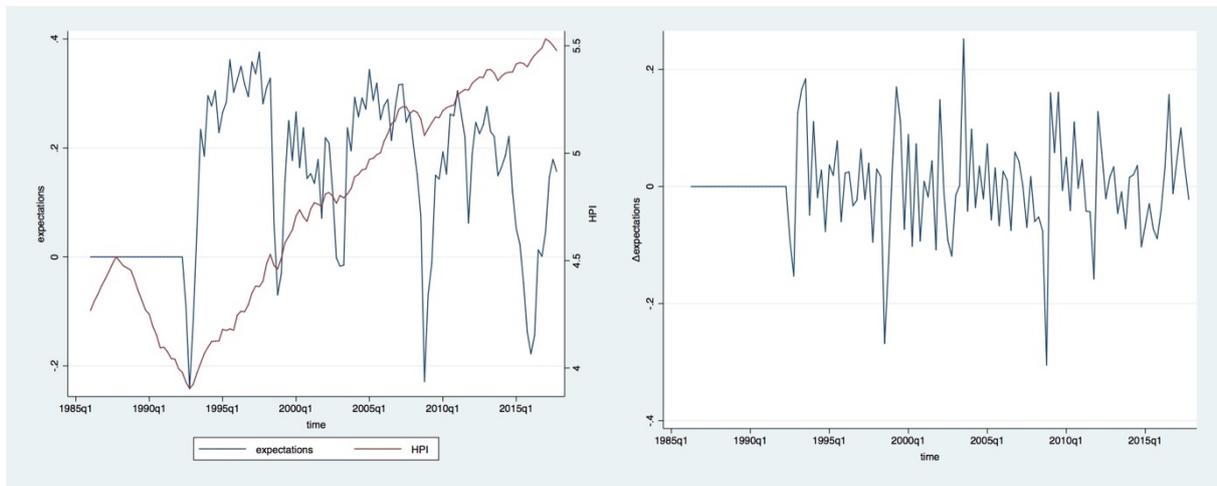
Figure 22: the households' expectations to the future economy and the differenced expectations.

6.8 The housing supply

6.8.1 The housing stock in fixed prices

From figure 23 exhibits how the housing stock is growing smoothly over time compared to the house price. It is not likely that the housing stock will return to a mean value. Since a house usually provides housing services for decades after it is constructed, I assume that a shock to the housing stock is permanent. Therefore I expect the variable to be non-stationary in levels. However, the graph over the differenced housing stock does not look clearly stationary either. From the Banking Crisis the housing stock is more or less constantly growing until the peak of the Financial Crisis. The unit root tests confirm that the variable is non-stationary in levels. Additionally, the unit root tests cannot reject the null of a unit root for the differenced data either. Unit root tests have weaknesses in finite samples, but by looking at the data in the graphs I conclude with the housing stock being I(2).[13]

---

[13] Anundsen and Jansen (2013) also found that the differenced housing stock (K83) variable was not stationary (from appendix C in their article)
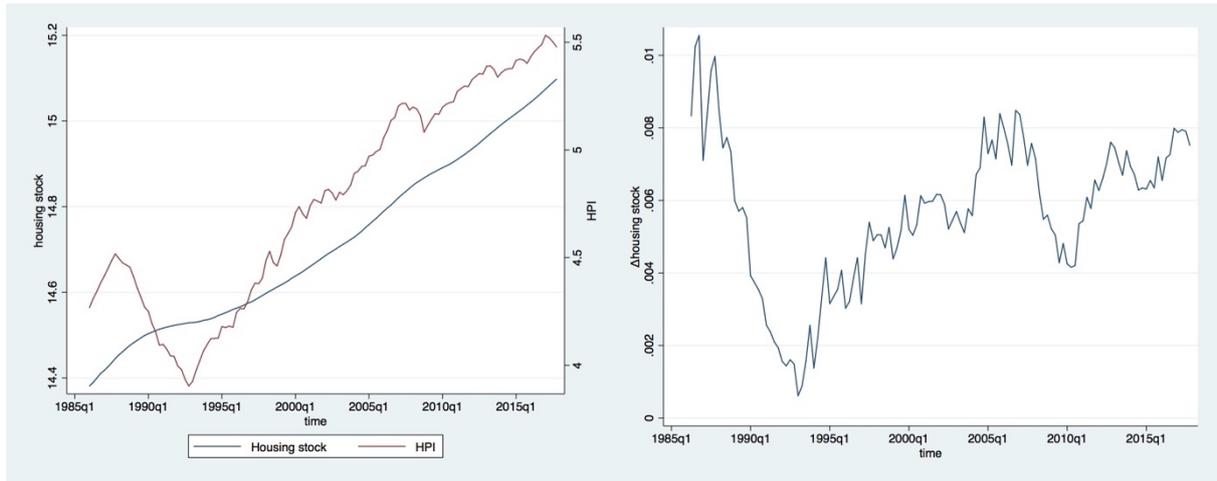
Figure 23: the housing stock measured in fixed prices

6.8.2 Initiated square meters of housing

Since the cointegration in error-correction models rely on variables that are of the same integration, the housing stock can be problematic. It is also an issue that the housing stock is I(2) for the ARDL model and the Bounds test, as it invalidates the model. Therefore, an alternative measure of the housing supply is included. As the initiated square meters of housing is supposed to be a proxy for the housing supply, the same arguments as for the housing stock applies. A house that is constructed yields services for many years. Therefore an increase in the amount of constructed houses would be a permanent addition to the housing stock. The graphs in figure 24 are somewhat in order with the expectations, besides the significant outlier at the beginning of the period. Both the ADF and the PP test concludes with stationarity both in level and in differenced form. A Zivot-Andrews test on the data in levels, cannot reject a unit root. Unit root test have weaknesses, and in this case the tests are being impacted by the extreme drop in construction following the crisis. The growth in the amount of initiated square meters have a positive average following the Banking Crisis. It is only momentarily negative following the Banking and the Financial Crisis. With a positive growth over time, the variable will not return to a mean value. Therefore the conclusion is that the initiated square meters of housing is integrated of order one I(1).
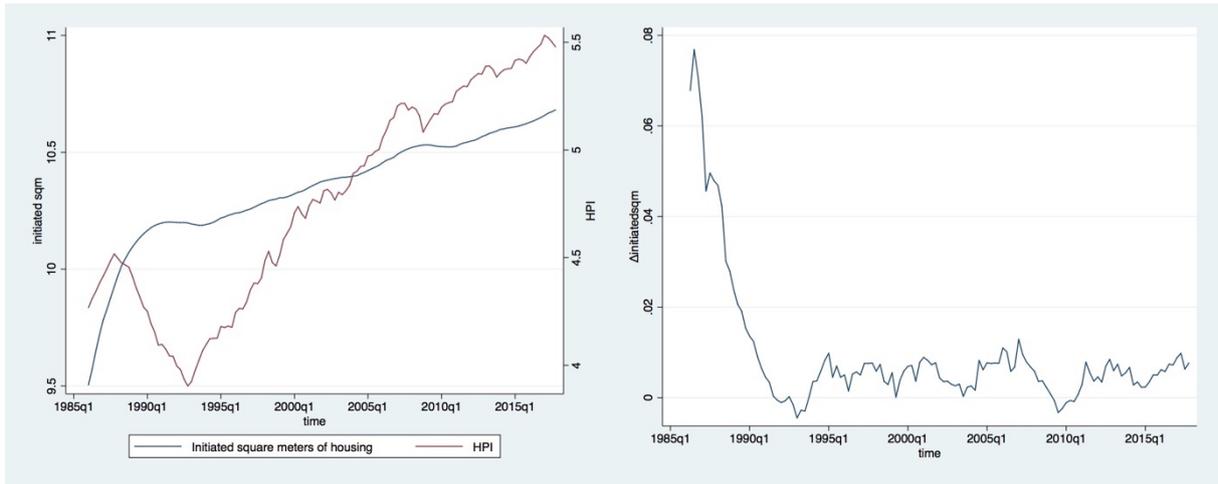
Figure 24: the total amount of initiated square meters of housing

7. Re-estimation of Jacobsen and Naug's house price model

This chapter re-estimates the house price model of Jacobsen and Naug (J&N) (2004). First the model is re-estimated over the original sample period, to compare the results, before it is estimated on a sample of almost twice the amount of observations from the original sample. It is important to notice that all results from the re-estimation of the J&N model is given by the available dataset. A re-estimation on the original dataset could potentially have given different results.[14]

In their paper J&N builds an empirical model for house price growth based on the one-step error-correction procedure. The motivation for the paper was that the house prices had tripled from 1992 until 2004. In addition to explain the house prices of the past, they also wanted to be able to predict the house prices in the near future. They use the model to examine whether the house prices were above a fundamental house price given by the explanatory variables in the model.

J&N emphasized a nominal version of their model. The reason for this was that they found consumer prices to have coefficients and t-values close to zero when they included the nominal interest rate as an explanatory variable. Additionally, models with the nominal interest rate showed a better fit than the model with the real interest rate. They attached the real version of their model in the appendix, which included the real house price, the real income and the real interest rate, deflated by CPI-ATE. This paper will focus on the model in real terms, and thus it is the real model of J&N that will be re-estimated and examined.

The variables considered by Jacobsen and Naug were the households' (nominal) income, indices for house rent paid and total house rent in the consumer price index (CPI), other parts

---

[14] Bjørn Naug at Norges Bank was very helpful with providing data. I received two datasets from Naug, where one was more or less identical to the original set they used to create their model in 2004. The other one was revised, updated and extended to the second quarter of 2017. However, the original income variable had been discontinued since they created the original model. The other variables were still similar to the original values, but with some revisions and modifications. Jørgen Ouren at SSB was able to provide me with an alternative series of the households' total wage income, and I assume it to be similar enough for a further analysis and comparison. From the comparison model it is evident that there are differences from the original dataset and the current one, and this is a potential weakness of the analysis.

of the CPI adjusted for tax changes and excluding energy products (CPI-ATE), various measures of the real after-tax interest rate, the housing stock (as measured in the national accounts), the unemployment rate (registered unemployment), backdated rise in house prices, household debt, the total population, the shares of the population aged 20-24 and 25-39, various measures of centralization and an indicator of the households' expectations to their own financial situation and the Norwegian economy. As J&N were considering a high number of potential explanatory variables compared to the number of observations, they estimated several alternative models where they only included parts of the variables, before they simplified the models by imposing restrictions that were not rejected by data.

House rents generally had coefficients and t-values close to zero. This may have reflected that rents in housing cooperatives accounted for an important share of house rent indices in the CPI during the estimation period. Additionally, many house rents were strongly regulated during the period and they did not have long time series of house rents in their dataset.

J&N found no significant effect of households' debt on house prices, even if they only included debt effects for the period 1990-1993. In isolation, the insignificance indicates that credit was not restricted by banks' profitability nor government regulations during the estimation period.

J&N tested for the banks' lending rates and various market rates. They found that the banks' lending rates were strongly significant in all models, while the market rates were clearly insignificant in models where the banks' lending rate was included. Potential reasons for this could be that the interest rate was used to stabilize the development of the exchange rate during the 1990s, and that the population used the observed interest as an estimate of future interest rates. The market rate captures to some extent the effects of a change in business cycles. Thus, both house prices and the difference between the banks' lending rate and the market rates were dependent on the economic outlook. It is therefore likely that the effect of interest rate expectations is undervalued in the estimated equations.

J&N found no evidence for a strong direct impact from population movements or demographic factors. They also found no significant effect of backdated growth in house prices, which according to their discussion is a sign of that households only to a limited degree use observed growth in house prices as an indicator of future growth in prices. If backdated growth is of little

importance for future price growth, then the danger of a housing bubble is reduced as there is not the same risk of a self-reinforcing price spiral.

The argument to include the unemployment rate was that increased unemployment gives expectations of lower income growth and increased uncertainty of the future purchasing power of oneself and others. This will in turn have an impact on the willingness to pay high prices for a dwelling. The unemployment can additionally be a measure of the business cycles. The unemployment rate is a variable that is a part of the preferred model.

J&N found that the indicator for the households' expectations to the economy was highly correlated with the development in house prices. Yet, it was also strongly correlated with the interest rate and the unemployment rate, which are also variables that capture the mood in the market. As a result, they adjusted the consumer confidence indicator for the effects of the interest rate and the unemployment rate. The estimation procedure and output can be seen in appendix C. The adjusted variable, EXPEC, represents the households expectations, less the effect of interest rate and unemployment rate. EXPEC are shifts that can be due to for example political conditions, or a change in the outlook for the Norwegian economy. The variable takes the value zero in any previous periods.[15]

J&N argue that the households' income and housing stock are strongly correlated when they are adjusted for seasonality, and as a result the effects of the variables are imprecisely estimated if they are included with separate coefficients. They cannot reject a hypothesis that the two variables have the same coefficient with opposite signs. Therefore they impose a restriction that the income and the housing stock should enter the model with the same coefficient but with opposite signs. In the real model, they impose the condition that the housing stock should be ¾ of the coefficient, so that the housing stock will have approximately the same long-term coefficient in both the nominal and the real model.

J&N do not mention in the article why they chose to start in 1990 and not before. In their paper they graph several of the variables back to 1986, such as the interest rate, the unemployment rate and the housing stock. In the original paper J&N normalizes the coefficients in the error-

---

[15] A unit root test of the adjusted households' expectations found no evidence of a unit root. Thus, the series are I(0).

correction term on the lagged house price. For comparison purposes the long-run coefficients in the re-estimated model has also been normalized on the house price, by dividing the original coefficient with the coefficient of the lagged house price. The preferred model of J&N, a comparison model on the current dataset and the extended model can be seen in table 1 below. The long-run variables are in italic letters.

Table 1: The original and re-estimated Jacobsen and Naug house price model.

Δhouseprice$_t$	Original J&N 1990-2004	Comparison 1990-2004	Re-estimation 1986-2014
Δincome$_t$	0.22***	0.641**	0.254*
	(3.37)	(2.01)	(1.84)
ΔINTEREST$(1-\tau)_t$	-3.10***	-2.836***	-2.102***
	(6.84)	(-4.90)	(-3.83)
ΔINTEREST$(1-\tau)_{t-1}$	-1.38***	-1.789***	-0.297
	(2.91)	(-2.87)	(-0.52)
EXPEC$_t$	0.05***	0.065**	0.062***
	(3.46)	(2.09)	(3.76)
houseprice$_{t-1}$	-0.17***	-0.128***	-0.087***
	(7.43)	(-4.79)	(-4.67)
INTEREST$(1-\tau)_{t-1}$	4.19***	3.14	15.05***
	(3.31)	(-1.34)	(-5.35)
unemployment$_t$	0.23**	0.078	0.23**
	(2.49)	(-0.61)	(-2.54)
(income- 0.75housingstock)	-2.26***	-2.78***	-1.21**
	(12.01)	(4.20)	(2.49)
Q1	0.02**	0.04***	0.0338***
	(2.10)	(5.83)	(5.75)
Q2	0.01	0.004	0.0096
	(1.35)	(0.34)	(1.54)
Q3	0.01	-0.0187	-0.0003
	(1.15)	(-1.51)	(-0.05)
Constant	-0.21***	0.213***	0.27***
	(5.67)	(2.82)	(4.92)
Observations	56	56	115
R^2	0.87	0.83	0.614

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Most of the coefficients in the extended model have similar magnitudes to the original model. However, the variable that clearly differs the most is the interest rate. This might be due to that

the sample contains two of the largest crisis in the modern Norwegian economy, while J&N only included parts of the Banking Crisis.

The coefficient of the lagged house price, which is also the adjustment speed, is highly significant. However, the magnitude has been slightly reduced. If house prices in the extended model rise 1 percent above their long-term value in quarter $t-1$, then the house prices will fall by -0.087 percent in quarter t. This is lower than the adjustment speed of -0.17 in J&Ns estimation. According to the re-estimated model, it will take $\frac{1}{0.087} \approx 12$ quarters, or three years to return to the equilibrium, after a 1 percent deviation from the equilibrium house price.

7.1 Long-run effects

The long-term effect of a 1 percentage point increase in the interest rate is a fall in house prices of 4 percent in the original model and 15 percent fall in the extended model. The results are clearly different, and the extended model has a much larger interest rate effect that the original model. The interest rate is clearly a very important factor to explain house price growth.

Unemployment in the re-estimated model has not only a similar coefficient to the one in the original mode, but it also has a similar t-value. J&N included the contemporaneous rather than the preceding unemployment rate in the long-run equation. Usually the variables in the long-run equation should be of the same period. A potential reason why they used the contemporaneous unemployment rate is the typical lag of the unemployment rate to other variables, such as the interest rate and the house prices. If the unemployment increases from 4 to 5 percent, then the house prices will fall with 5.75 percent in the long-run according to the re-estimated model.

A 1 percent increase in the income will increase the house prices with 1.2 percent in the long-run, while a 1 percent increase in the housing stock will increase the housing supply and thereby reduce the house price with 0.9 percent.

7.2 Short-run effects

According to the extended model a one percent increase in the households wage income increase house price growth with 0.25 percent. The coefficient is stable as it is close to the

original model that implies a 0.22 percent increase in house price growth from a 1 percent increase in disposable income.

A 1 percentage point increase in the interest rate implies a 2.1 percent decrease in the growth of house prices, compared to a 3.1 percent decrease in the original model. The lagged differenced interest rate has a much lower magnitude, and it is insignificant in the extended model. The coefficients of the modified household expectations are very stable, as both the magnitude and the t-values are similar for both the original and the extended model. The effect implies that house prices are primarily affected by large shocks to the expectations. Small changes in the expectations can represent noise in the data, as it is based on a sample survey.

7.3 Examination of residuals

The extended model contains almost twice as many observations as the original model. While J&N included only the Durbin-Watson statistic as a measure of serial correlation, the re-estimated model will be tested for serial correlation with the Ljung-Box test and the Breusch-Godfrey test. The two latter tests are more precise when it comes to determining the degree of serial correlation in a model, as the Durbin-Watson test have indecisive areas where we cannot conclude.

Table 2: Post-estimation tests of the re-estimated Jacobsen and Naug model

	Original J&N 1990-2004	Comparison 1990-2004	Re-estimation 1986-2014
Observations	56	56	115
Serial correlation:			
Durbin-Watson test	2.47	2.43	1.41
Ljung-Box test	-	3.34*	7.63***
Breusch-Godfrey test	-	3.18*	10.82***
Heteroscedasticity:			
Breusch-Pagan test	-	17.6*	11.2
Normality:			
Skewness/Kurtosis test	-	7.45**	0.65

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

The model from 2004 does not have a particular issue with serial correlation, as the null hypothesis of no serial correlation is only rejected at the 10% level. For the full model, however, there are strong evidence of serial correlation, as both the Ljung-Box and the Breusch-Godfrey tests reject the null hypothesis of no serial correlation at the 1% level. The Durbin-Watson statistic is slightly below the lower critical value and thereby supports the conclusion of serial correlation. Serial correlation in the residuals violates the ordinary least squares assumption that the error terms are uncorrelated. This means that the Gauss Markov theorem does not apply, and OLS estimators are no longer the Best Linear Unbiased Estimators (BLUE). While it does not bias the OLS coefficient estimates, the standard errors tend to be underestimated and the t-scores overestimated, when the serial correlation of the residuals at low lags are positive. As a result, coefficients can be wrongfully accepted as significant.

J&N did not include any information about the distribution or heteroscedasticity of their residuals. Figure 25 shows the residuals from the estimation of the J&N model on the full sample, while figure 26 shows the distribution of the residuals. It does not look like we have an issue with heteroscedasticity, which is confirmed by the Breusch-Pagan test for heteroscedasticity in table 2.
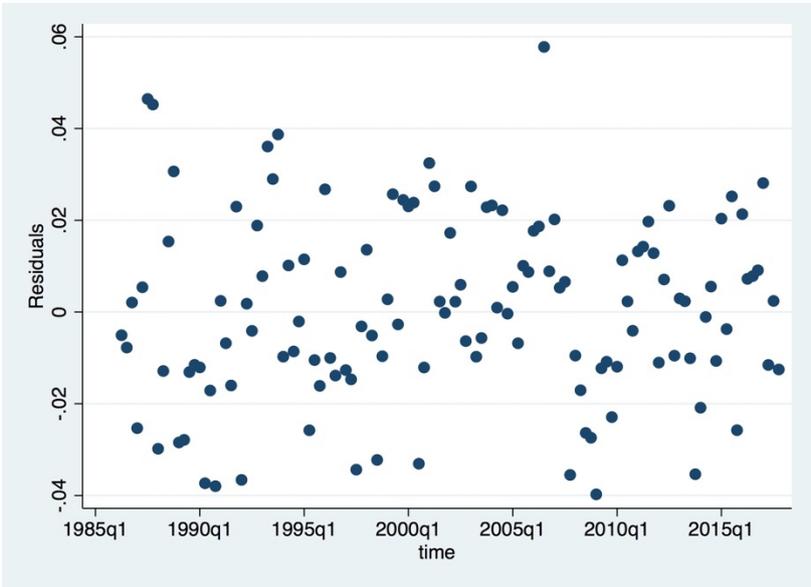
Figure 25: the residuals from the full J&N model

The distribution of the residuals do not look too normally distributed, especially for the model on the original sample. However, the Skewness and Kurtosis test in table 2 does not indicate any big problems with the normality of the residuals in the extended model.
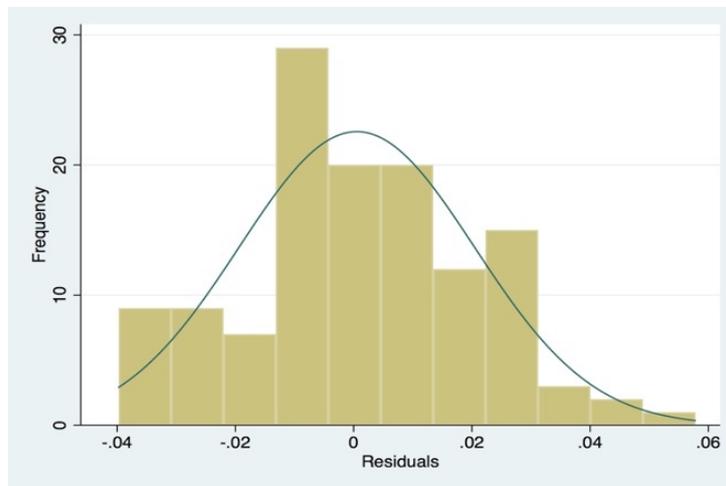
Figure 26: distribution of the residuals from the J&N model

EViews offer CUMSUM tests of the residuals and the squared residuals to check the stability of the estimated models. The CUMSUM test is based on the cumulative sum of the recursive residuals (Brown, Durbin, & Evans, 1975). The implemented test in Eviews 10 plots the cumulative sum together with the 5% critical lines. If the cumulative sum goes outside the critical lines, it is an sign of parameter instability. The cumulative sum in figure 27 stays well inside the 5% critical lines, and the model looks stable.
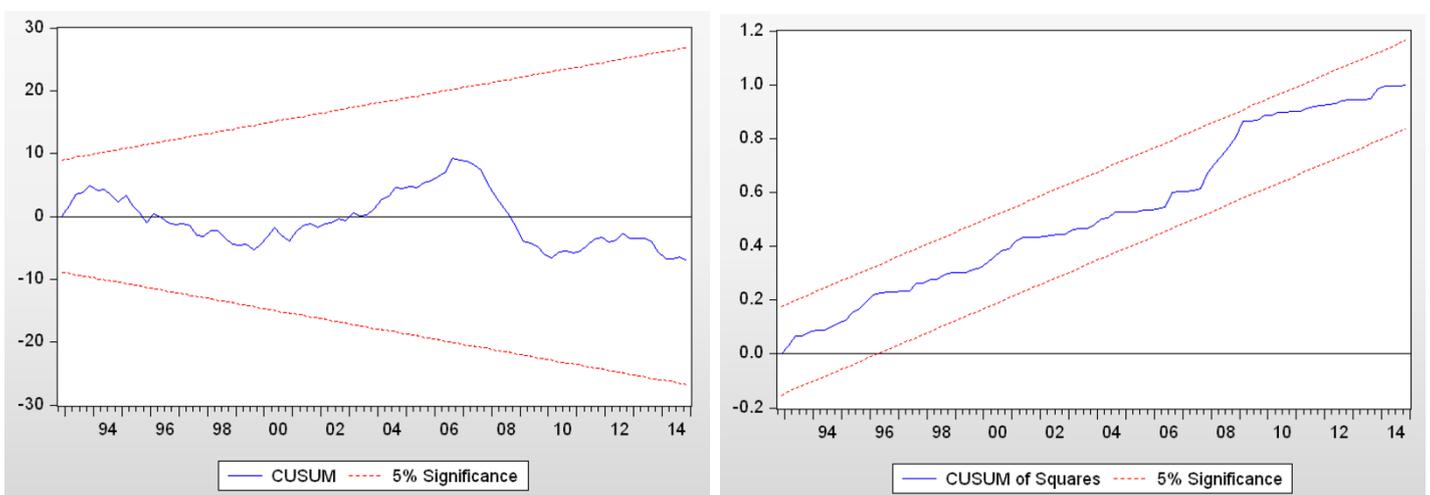
Figure 27: the CUMSUM test and CUMSUM of squared residuals test for the J&N model

7.4 In-sample fit

Static prediction uses the actual house price and the actual explanatory variables when estimating one step ahead. Dynamic prediction on the other hand, utilizes the estimated house price from the last period and the actual explanatory factors. As a result, potential deviations of the predicted values from the actual house price will be transferred to the next period, as the actual house price is not taken into account, to recover the former deviation. Thus, the dynamic prediction is a "tougher" in-sample fit test, than the static prediction. All predictions in the rest of the paper will be dynamically estimated.
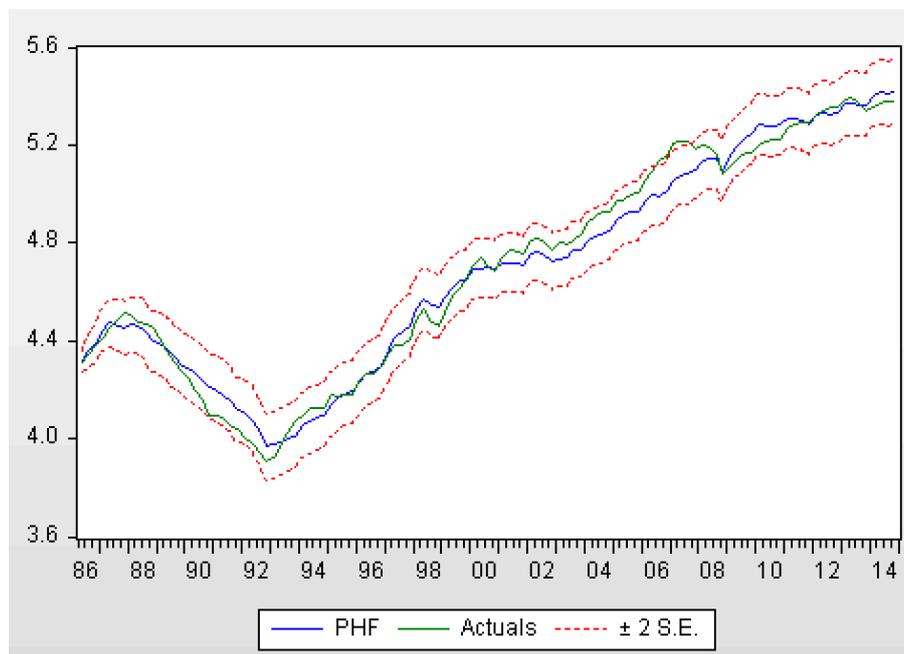
Figure 28: dynamic in-sample prediction of the J&N model.

In-sample prediction[16]	RMSE	MAE	MAPE
J&N model	0.06	0.049	1.05

The actual house price goes outside the two standard deviation bands around the Financial Crisis, and it is close to go beyond the bands during the drop in house prices following the Banking Crisis. In addition, it underpredicts the house price more or less constantly from 2000 until 2008. The in-sample forecasting tests do not provide much information in isolation, but can be helpful to compare the fit to the data against the alternative model.

---

[16] See (Brooks & Tsolacos, 2010) for a definition of the forecasting evaluation tests.

7.5 Potential issues with the model

First of all, the J&N model does not seem robust, as some of the coefficients changed dramatically when I extended the sample. Especially the effect of the interest rate was clearly different from the original model. In addition the residuals are serial correlated according to both the Ljung-Box and the Breusch-Godfrey tests. To cope with serial correlation, it is normal to include more dynamics in the model, by allowing more lagged variables in the short-run.

In the article, J&N provides no information of the variables stationarity. The only indication is the direct quote "*If the housing stock and wage income grow at the pace prevailing for the last five years, house prices will increase by about 5 per cent per year for given values of the interest rate, the unemployment rate and the (adjusted) consumer confidence indicator. Since these variables are stationary, this means that house prices will rise in pace with wage income in the long term*" from page 7 in the English version of the article. Thus, J&N assume the nominal interest rate, the unemployment rate and the expectation indicator to be stationary. In chapter 6 the unit root analysis found the interest rate to be I(1), the unemployment rate to be I(0). In addition, I concluded with the housing stock being I(2). However, unit root tests have often low power in finite samples, so the conclusions are not always robust.

For ordinary cointegration tests to make any sense, all variables must have the same order of integration. In the one-step method it is not common to test the estimated residual of the long-term values for stationarity like in the Engle-Granger two-step method. The significance of the error-correction term (the lagged dependent variable) is used to test for cointegration. In both the original model and the extended model the error-correction term is significant and thus indicates cointegration. Given that the long-run equation contains variables with different order of integration, it is unclear whether there is a cointegration between the long-run variables in the J&N model. An alternative model should be properly tested for a cointegration between the long-run variables. In addition, the model should allow for more dynamics in order to deal with the serial correlation in the residuals found by the serial correlation tests. In the following chapter, an alternative model with these characteristics is estimated.

8. An updated single-equation model for Norwegian house prices

This chapter estimates an ARDL model for Norwegian house prices following the methodology suggested by Pesaran et al. (1999). An ARDL model allows for both short-run dynamics and long-run relationships, like the traditional error-correction models. The re-estimated version of J&N's model had clear issues with serial-correlation. By including more dynamics in the model it is possible clean up the residual serial correlation. If an ARDL model has well behaved residuals with no serial correlation and heteroscedasticity, the Bounds test can be used to test for cointegration between the long-run variables.

The unit root analysis clearly concluded with the housing stock being I(2). Including an I(2) variable in the ARDL model will invalidate the methodology. Therefore the initiated square meters of housing will in the following be used to represent the supply side in the model.[17] As the income variable, and the housing supply variable no longer has a particularly strong correlation, the variables will enter the long-run equation individually.

The analysis suffers from bad data when measuring population effects on the house price. While the frequency of the dataset is quarterly, most demographic data of the Norwegian population is published annually. The only data on population characteristics that is published quarterly by SSB is the total population growth. However, the series only go back to the fourth quarter of 1997, and it has little variation in levels. The annual series before 1997 was extended by linear disaggregation, but as expected the series provided little additional explanation of the house price. Demographic changes are slow moving, and especially on a small sample it can be hard to identify an effect. Yet, a growing population implies a larger workforce. This will increase aggregated level of income for a given wage. When the population increases there will in addition be a larger demand for housing. As a result, people are forced to spend more of their income on housing. Thus, an increasing population will be reflected to a certain extent in the aggregated wage income, and some of the income effect on house prices is most likely coming from demographic changes. Direct demographic effects could not be identified, however, and was therefore left out of the preferred model.

---

[17] By re-estimating the J&N model with initiated sqm of housing, it was clear that it did not change the conclusion of serial correlation, which was expected as the housing supply only enter the equation in the long-run.

The increasingly high debt levels of Norwegian households can be a potential source for financial instability. As a result, the Norwegian FSA introduced credit regulations to limit the growth in household debt. If debt is limited by credit regulations from the government, there can be credit effects on the house price that the interest rate is not picking up. Since 2010, there has been introduced several new regulations towards mortgages. Therefore it would be interesting to see if this would impact the significance of the households' debt in the model. In addition to government regulations, the banks' profitability can limit the households' debt. If the profitability of the bank goes down, it is less inclined to provide an increasing amount of loans. Thereby the growth in credit will go down. In the estimation period, there are several periods where the banks' profitability have been reduced, such as the Banking Crisis in the late 1980s, and during the Financial Crisis in 2008. J&N found no significant effect of debt in their sample.[18] Therefore, debt will be included as an explanatory factor to see if the new regulations or the banks' profitability has had a significant effect of debt on the house prices. When I perform an omitted variable test of debt on the initial ARDL model, it rejects the null that debt is not significant. Thus, the test supports my expectations that debt should be included in the model.

The unemployment rate is clearly insignificant both in the short- and long-run. A F-test and a likelihood-ratio test supported the removal of the unemployment rate. As a result, the unemployment was left out of the model. House rents are clearly insignificant when it is included in the model, which was also what J&N found during their estimations. Thereby it is left out of the preferred model, as it added little value.

First a trend restricted to the long-run equation is included, as both the real house price, the real income and the initiated sqm of housing have an up trending behavior. However, the trend is insignificant. Following Pesaran et al. (1999), the statistically insignificant trend is left out of the model. The included constant in the ARDL model is also clearly insignificant. Removing the constant has a minimal impact on the coefficients, and therefore the constant is left out of the model.

---

[18] Anundsen and Jansen found in 2013 evidence for that debt has an important impact on house prices multi-equation model. In addition, they found that there exists a self-reinforcing effect between house prices and household debt (Anundsen and Jansen, 2013)

When the model is constrained to have maximum four lags for both the dependent and the independent variables, the software includes a minimum of lags. As a result, the model suffers from heteroscedasticity and instability. Accordingly, the maximum lag length is increased to five, and the ordinary coefficient covariance matrix is changed to a Huber-White-Hinkley covariance matrix for robust standard errors. The estimated long-run relationships are presented below.

8.1 The long-run relationships

The equations have been normalized on the house price. The first model includes the real house price, the households' real income, the real interest rate and the initiated square meters of housing in the long-run equation. The second model includes the same variables, in addition to the real debt of the households. All the long-run coefficients in both models are clearly significant at the 1% level, and all have the expected signs.

Model 1 - without debt

EC^{19} = houseprice – (1.5 income – 1.3 initiatedsqm – 12.4 INTEREST)

Model 2 – with debt

EC = houseprice – (1.5 income – 1.8 initiatedsqm – 6.4 INTEREST + 0.4 debt)

An increase of 1 percent in the households' income will in the long-run increase house prices with 1.5 percent according to both models. The coefficients are not very different from the re-estimated J&N model, where the effect of a 1 percent increase in the income would lead to a 1.2 percent increase in the house prices.

A 1 percent increase in the housing supply will in the long-run reduce house prices with 1.34 percent, according to model 1, and 1.81 percent according to model 2. The corresponding effect in the re-estimated J&N model was 0.9 percent. Thus, the supply effect in the ARDL model with debt is almost double the effect from the re-estimated J&N model.

---

[19] EC = error-correction term

A 1 percent increase in debt, will only increase house prices with 0.36 percent in the long-run. As there was no significant effect of unemployment in the model, the effect of the debt on house prices must come from a constraint on the banks' profitability or government regulations.

The re-estimated J&N model found that house prices would fall 15 percent from a 1 percentage point increase in the interest rate in the long-run. If the interest rate in the ARDL models rise permanently with 1 percentage point, model 1 estimates a fall in house prices of 12.4 percent compared to a fall of 6.6 percent in model 2. This is an interesting finding. A potential reason for the large interest effect is that the sample includes both the full period of the Banking Crisis and the Financial Crisis. In the 1990s the interest rate was used as a tool to stabilize the short-run exchange rate of the Norwegian currency (Jacobsen & Naug, What drives house prices?, 2004). As a consequence, the interest rate did not adjust rapidly to the developments in the house prices. From 1987 until 1993, the interest was continuously rising towards historically high levels, while the real house prices were falling. Either the real interest rate is very important in predicting the real house price, or the J&N model and the ARDL model without debt is putting too much weight on the interest rate.

8.2 Examination of residuals

Table 3: testing the ARDL models' residuals

	ARDL Model 1	ARDL with debt Model 2
Observations	111	111
Serial correlation:		
Ljung-Box test	0.0036	0.0069
Breusch-Godfrey test	0.339	0.199
Heteroscedasticity:		
Breusch-Pagan test	22.13[*]	20.75
Normality:		
Skewness/Kurtosis test	3.08	1.83

[*] $p < 0.1$, [**] $p < 0.05$, [***] $p < 0.01$

Before the Bounds test can be used to check for cointegration, it is essential that the residuals are well behaved. According to the residual tests in table 3, there is no evidence of serial correlation, heteroscedasticity or non-normality.

Figure 29 and 30 exhibits the CUMSUM and the CUMSUM of squared residuals tests for model 1 and 2 respectively. As before, if the cumulative sum of residuals move outside the 5% bands, it is a sign of instability. The cumulative sum is going marginally outside the 5% critical line for the ordinary CUMSUM test of model 1. The CUMSUM test of the squared residuals are stable, and the conclusion is that the models are stable. However, according to Pesaran et al. (1999), the CUMSUM test is known to have low power, so the results should be interpreted with care.
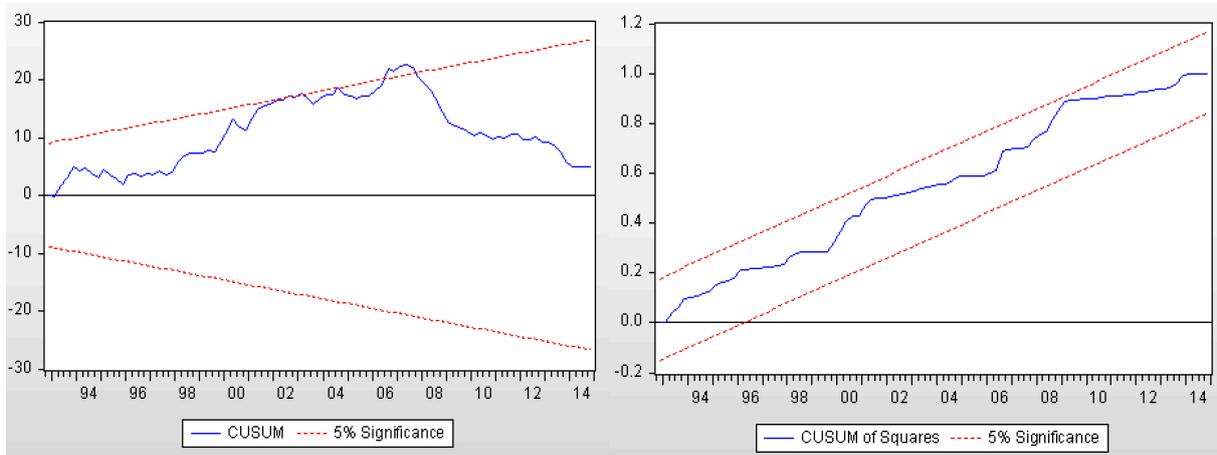

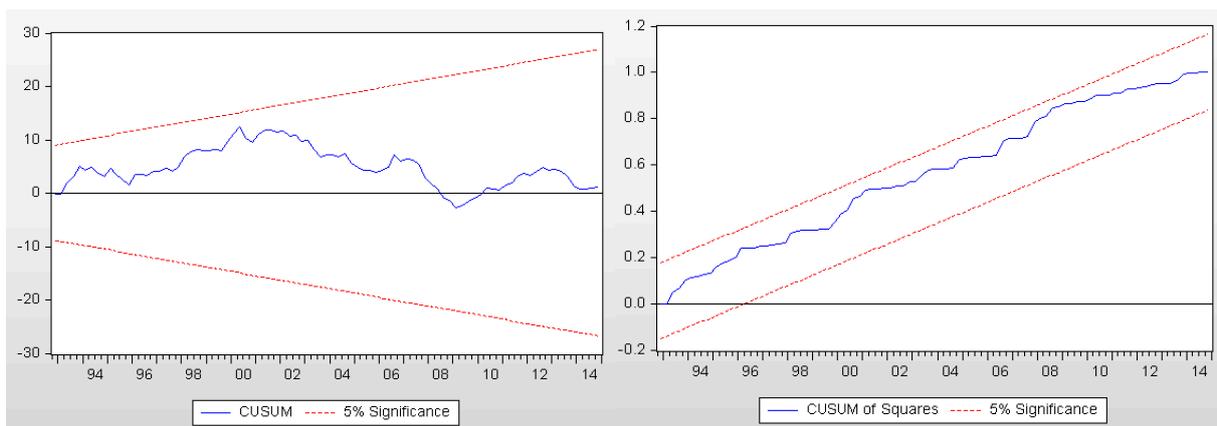*Figure 29: the CUMSUM and squared CUMSUM test for model 1*

Figure 30: the CUMSUM and squared CUMSUM test for model 2

8.3 Bounds testing

Table 4: The Bounds test of the ARDL models[20]

F-test	Test statistic[21]	5% lower bound	5% upper bound
Model 1	5.78	2.45	3.63
Model 2	6.94	2.26	3.48
t-test			
Model 1	-4.50	-1.95	-3.33
Model 2	-5.48	-1.95	-3.6

It is essential to confirm the existence of a long-run level relationship between the long-run variables for the ARDL model to be valid. There are no issues with the residuals, which makes it possible to utilize the Bounds test on the long-run relationships. As in conventional cointegration testing, the Bounds test is testing for the absence of a long-run relationship. A rejection of the null implies that we have a long-run relationship. All the included variables in the long-run equation are assumed to be I(1) from the unit root analysis. Both the F-test and the t-test statistics in table 4 are above the upper 5% critical bounds, and thus confirm that there are

---

[20] If the data has more than one potential long-run relationship, the ARDL model is inappropriate. The household income is most likely not impacted by the house prices besides the general development in the economy. The interest rate is utilized to adjust inflation and control the exchange rate, and thus it has other purposes than just controlling the price on mortgages. For instance, the interest rate has been low during the last years, even though the house prices have been rising. The housing stock is being impacted by the house prices, but in a very slow manner, and therefore the housing stock is given as exogenous in the analysis. The household debt is the variable that is of more concern. Anundsen and Jansen (2013) found two cointegrations between the real house price, the real income and the real debt. They normalize the two cointegrations on debt and house prices, and let several weakly exogenous variables enter the cointegrating equations. Anundsen and Jansen do not include the interest rate in the error-correction term of housing, and explain that according to their assumptions, the interest rate works through the debt error-correction term in the long-run. One potential implication of only having one cointegration, is that the interest rate will have a too large effect, and the debt a too small effect. However, a Johansen's test on the current dataset identifies only one cointegrating relationship between the variables in both models. The result can be seen in appendix D.1. Statistical tests can often be inaccurate, but there is only proof of one cointegration between the variables, and thus it is assumed that there exists only one potential cointegration between the variables in the further analysis.

[21] All four test statistics are even significant at the 1% level.

long-run relationships, or cointegration, in both models. However, since the Bounds test confirms that there exists a long-run equation when debt is included, that implies that model 1 is misspecified. But for now, both models will be considered in the further analysis.

8.4 The short-run dynamics

Table 5: Short-run coefficients of the ARDL models

Sample: 1986Q1-2014Q4	ARDL Model 1	ARDL with debt Model 2
Dependent variable	$\Delta houseprice_t$	$\Delta houseprice_t$
EC_{t-1}	-0.065***	-0.101***
	(-4.50)	(-5.48)
$\Delta houseprice_{t-1}$	0.33***	0.28***
	(4.01)	(3.34)
$\Delta houseprice_{t-2}$	0.021	-0.025
	(0.25)	(-0.31)
$\Delta houseprice_{t-3}$	-0.12	-0.11
	(-1.41)	(-1.31)
$\Delta houseprice_{t-4}$	0.29***	0.24***
	(3.83)	(3.20)
$\Delta INTEREST(1-\tau)_t$	-1.49***	-1.49***
	(-2.94)	(-3.06)
$\Delta debt_t$		0.18*
		(1.78)
$EXPEC_t$	0.049***	0.053***
	(3.25)	(3.65)
q1	0.023***	0.025***
	(3.53)	(3.96)
q2	-0.007	-0.0023
	(-1.00)	(-0.33)
q3	-0.007	-0.0048
	(-1.05)	(-0.75)
Observations	111	111
R^2	0.73	0.75
Lag selection	AIC	AIC

t statistics in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

As we have confirmed a cointegration between the long-run variables, we can include the error-correction term in the short-run dynamics.[22] If the real house price in model 1 rise 1 percent into disequilibrium, the real house price will fall - 0.066 percent in the next period, which is not very different from the adjustment speed of - 0.087 in the extended J&N model. It would take $\frac{1}{0.066} \approx 15 \ quarters$ to return to equilibrium. A potential movement of 1 percent into disequilibrium for the house price in model 2 will be corrected - 0.1 percent within one period. It would take $\frac{1}{0.1} = 10 \ quarters$ to return to equilibrium.

The lagged growth in the interest rate is included in both models, while the lagged growth of debt is included in model 2. Both models only includes the housing supply (initiated square meters) in the long-run. This is in accordance with economic theory, as we assume the supply of housing is given in the short-run. The more surprising feature is that both models do not include the growth of the households' income in the short-run, which differs from the re-estimated J&N model. Anundsen and Jansen (2013) on the other hand, do not include income growth in the short-run dynamics of their house price equation. According to the model, income is only having an impact on the house prices in the long-run. It is unlikely that the income would rise dramatically in the short-run, and especially enough to cover the expenses of a new dwelling. An increase in the households' income will increase their expectations about their future economy. Households are more inclined to start to saving money, in order to purchase a dwelling in the future. As such, the growth in income will increase their purchasing power in the long-run, but not have an immediate effect on the house prices.

The models include four lagged differenced terms of the real house price. Since the models have no issues with serial correlation, it is likely that the lack of autoregressive dynamics was an important reason for the serial correlation in the re-estimated J&N model. J&N found no effect of backdated growth on house prices, while the effect of backdated growth is clearly an important part of the estimated ARDL models. IMF found in 2004 that backdated growth in house prices had strong effects in the house price equation for 18 OECD countries, while Anundsen and Jansen (2013) found that the change growth one year ago impacts the house price. In the ARDL model, especially the preceding quarter's growth and growth one year ago impact the house price as they are significant and have reasonable coefficients. The other two

---

[22] Graphs of the cointegrating relationships can be seen in appendix D.2

backdated house price growth variables are not significant, and have small coefficients. According to the ARDL models a 1 percent growth in house prices one year ago leads to a $\approx$ 0.3 percent growth in the current period, which is close to the 0.38 percent effect found by Anundsen and Jansen in their house price equation. A 1 percent growth in the preceding quarter is also leading to a $\approx$ 0.3 percent growth in the house price, which was not found by any of the former studies.

The ARDL models find that the backdated growth is more important in the immediate short-run than Anundsen and Jansen. When backdated growth is important, households use observed growth in house prices as an indicator of future growth in house prices. This increase the danger of a housing bubble as there is a risk of a self-reinforcing house price spiral. It is economically plausible that the house price growth in the previous quarter and one year ago is important in explaining the house price development in the short-run. The price development the preceding quarter is an indication of the immediate expectations to a change in house prices. This can lead to impulsive purchases as households, who for instance have been saving money from an increase in their income, can expedite their house purchasing plans. As such, the effect can intensify the house price growth in a short amount of time. The one-year change in the house price is an indication of the general development in the house price, and it can make people more inclined to invest in the housing market as they perceive the market to be growing over time.

8.5 In-sample fit

In the following, the ARDL models' dynamic fit will be presented.
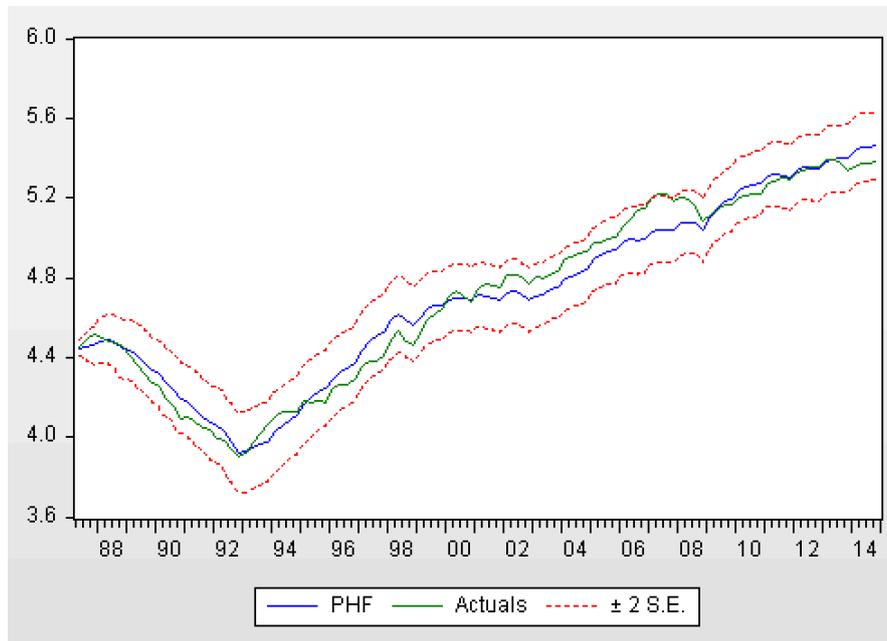
Figure 31: in-sample dynamic prediction for model 1 – without debt
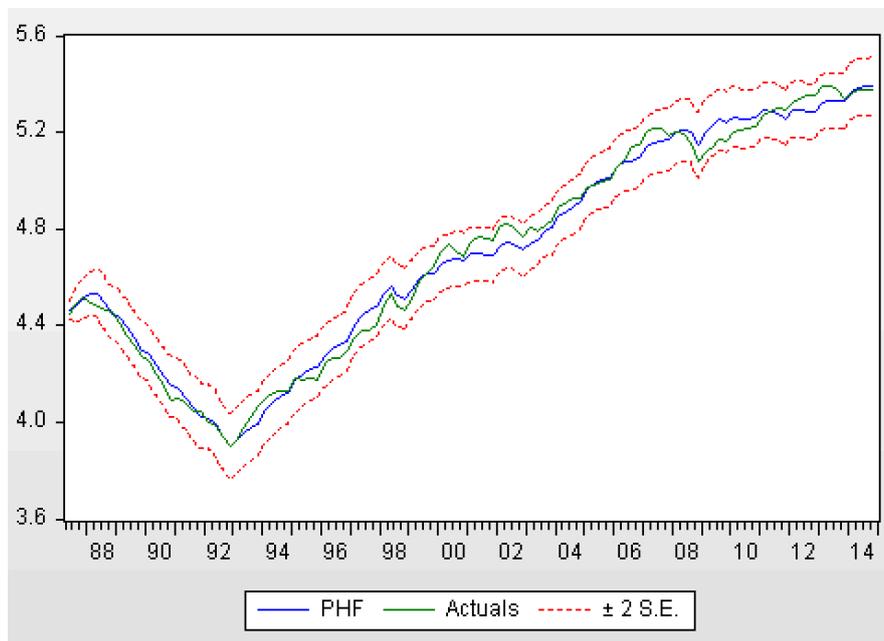
Figure 32: in-sample dynamic prediction for model 2 – with debt

The prediction of model 1 resembles the in-sample prediction of the re-estimated J&N model. The actual price is marginally inside the standard deviation bands around the Financial Crisis. In addition, the model underpredicts the house price consequently from 2000 to 2008. From 2008 it is able to follow the house price much more accurately. Model 2 on the other side,

performs much better. There are no periods where the actual price goes outside the standard deviation bands. But more importantly, model 2 is able to align much faster with the actual house price in the period between 2000 and 2008, most likely since the debt levels were growing from 2000 until 2008. Thus, the households' debt were an important factor to explain the house price development during the period leading up to Financial Crisis.

| In-sample accuracy	RMSE	MAE	MAPE
Model 1	0.07	0.06	1.30
Model 2 (with debt)	0.04	0.037	0.78

From the overview of the in-sample accuracy above, where a lower number is better, model 2 has a much better in-sample dynamic fit. This is another indication of that the model without debt is misspecified. Yet, the purpose of the model is to be able to forecast the house prices in the future. It is a bad idea to choose a model based upon the in-sample accuracy, as it might be overfitted to the data. Next, the models will be used to forecast out of sample, which will give a good indication of the models' robustness.

8.6 Forecasting

The forecasting is performed with actual explanatory variables and simulated house prices, which is also called dynamic prediction. The coefficients of the model are "locked" from the end of the estimation period. It is not a completely "true" forecast, which would predict the house price without the updated explanatory variables, but it is the same forecasting technique as used by J&N and Anundsen and Jansen. The out-of-sample time frame is 2015Q1-2017Q4, equal to 12 quarters.

In order to evaluate the forecasting properties of the estimated models, it can be useful to compare the forecasting results to a naïve model. The naïve model used in this case is a simple random walk, where $\rho = 1$ so that the best forecast of next period's value is the last period's value:

$$\Delta y_t = \Delta y_{t-1} + u_t$$

Hence, the forecast of the naïve model is completely flat throughout the whole forecasting period.
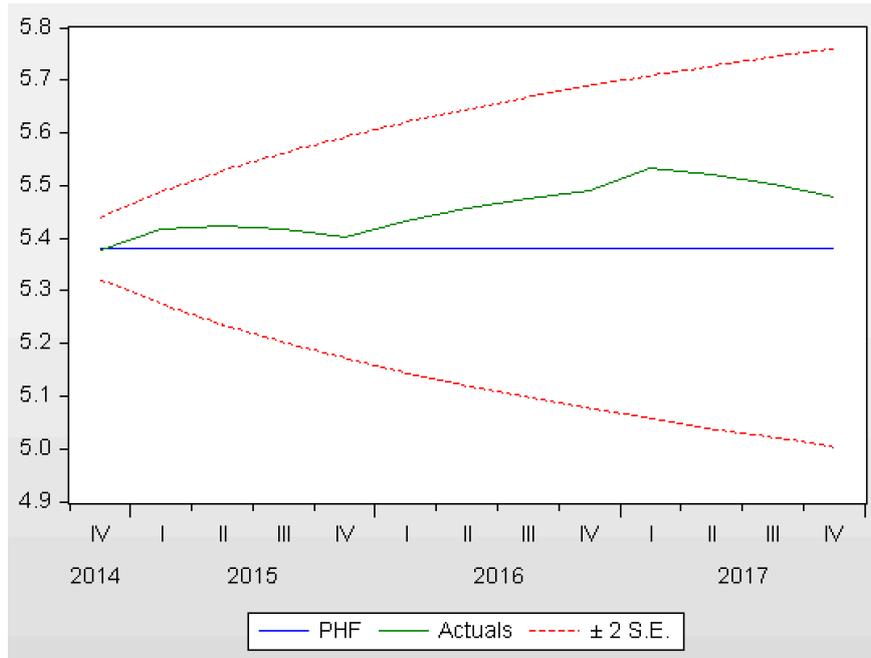
Figure 33: forecast of the naïve model

The ARDL model without debt forecasts the real house price very well. According to model 1, the real house prices were overvalued from the second quarter of 2016, and out through the sample.
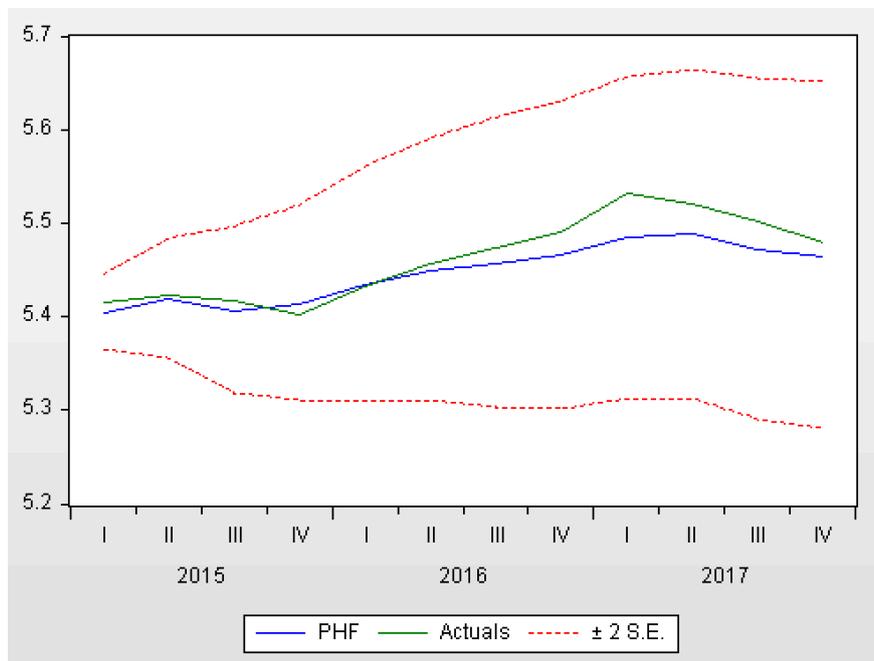
Figure 34: forecast of ARDL model 1 – without debt

The ARDL model with debt clearly underpredicts the house price in the whole forecasting sample. This was not according to the expectations, since the model that included debt in the

equation had a much better fit to the data in-sample. It is a good evidence of how a good in-sample fit never is a perfect indication of the forecasting properties of a model. The potential reasons for the bad forecast is that including the debt over fitted the model in-sample, and that the extreme change in total household debt during the Financial Crisis is impacting the stability of the coefficients.
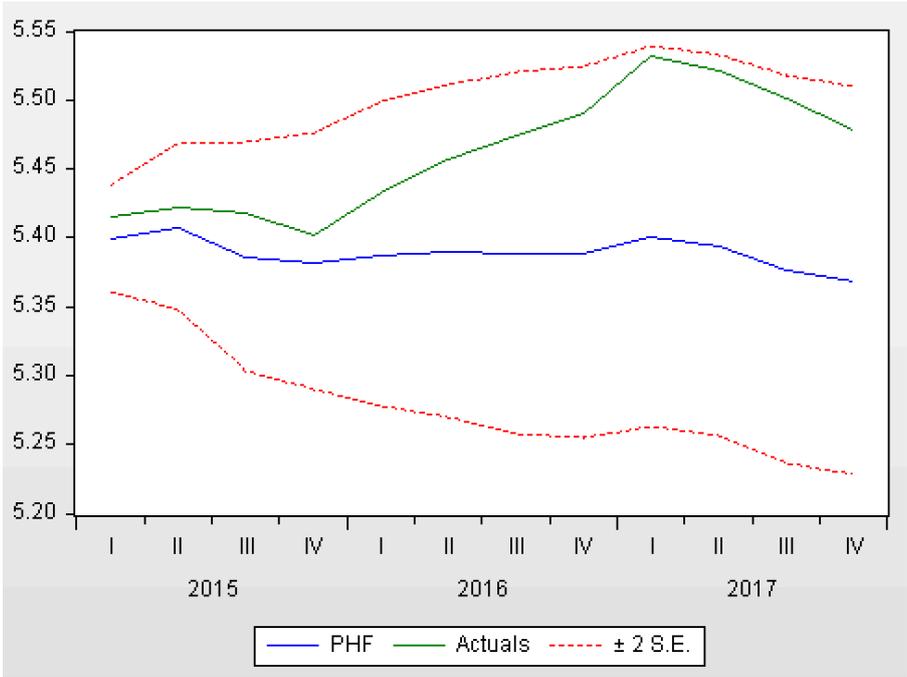
Figure 35: forecast of ARDL model 2 – with debt

8.7 Forecast performance

| Out-of-sample forecast	RMSE	MAE	MAPE
Naïve model	0.0875	0.075	1.36
ARDL without debt	0.022	0.017	0.32
ARDL with debt	0.085	0.073	1.33

The results above are a bit surprising. While the ARDL model with debt had a better fit in-sample, it is clearly beaten by the model without debt in the out-of-sample forecast.

The households' real debt kept increasing through the spring and summer of 2017, even though house prices were falling. As a result, the debt reduced the accuracy of the forecasting out of sample. Also, it is not unlikely that the development of debt around the Financial Crisis

generated instability in the model. The low interest rates were one of the potential reasons for why the housing market was heating up through 2016.[23] Households use the interest rate to create expectations of the future, and a lower interest rate makes it more attractive to purchase a dwelling. The result is higher activity in the housing market, and thus the house prices are expected to go up, as some households expedite their purchase due to the low price on credit. The interest rate was not expected to fall any further in 2017, thus most of the interest effect was likely taken out already in 2016 and early 2017. The effect of the interest rate on house prices in the long-run equation is much higher in the model without debt, and as a result the model is better able to forecast the out of sample house price. Therefore, the ARDL model without debt is the preferred model to describe the relationship between the house price and the fundamental drivers. Nevertheless, it is important to interpret the results with care. The model would for instance not be able to forecast as accurately if there was a large shock to debt that would impact the buyers' purchasing power, like during the Financial Crisis.

To summarize the analysis, the most important drivers of house prices in Norway are the real disposable income, the real interest rate, the housing supply and the households' expectations. In addition, the backdated growth in real house prices is an important factor in describing the short-run growth of the real house price. According to the estimated model with the fundamental drivers, the house prices in Norway were above their fundamental value from the second quarter of 2016 until the fourth quarter of 2017.

---

[23] The real interest rate was almost zero in the third quarter of 2016. A negative real interest rate implies that credit has no cost. You might actually "get paid" to take up credit.

9. Conclusion

This thesis has developed an econometric model to analyze the relationship between the real house price and the fundamental drivers of the housing market. According to the analysis the house price can clearly be explained through fundamental drivers of the housing market. The main drivers of the Norwegian house price are the households' real income, the real interest rate, the housing supply, backdated growth in house prices and the households' expectations to the future of the economy.

One of the most central house price models for Norwegian house prices the last decade, Jacobsen and Naug's model from 2004, had issues with serial correlation in the residuals after a re-estimation on a sample twice size of the original. To deal with the serial correlation an alternative ARDL model with more dynamics was estimated. The following analysis found no effect of house rents or population changes in the model. However, the data material on demographic changes was not ideal, but population effects are to some degree picked up through the households' accumulated income. The households' debt is significant and improves the in-sample performance of the model, especially around the Financial Crisis. In addition, by including the households' debt as an explanatory factor, the magnitude of the interest rate in the long-run relationship is reduced to a lower level. However, when an out-of-sample forecast is used to test the model's robustness, it becomes evident that including the debt overfits the model to the in-sample data. The preferred model does therefore not include the households' debt, but emphasize the real interest rate to a large extent.

In the out-of-sample period the real interest rate was close to zero, which made credit cheaper and increased expectations of future house price growth. As the preferred model emphasize the interest rate, it is able to accurately forecast the development of house prices in the out-of-sample period from 2015 to 2017. The interest rate was not expected to fall any further in 2017, and most of the interest effect was likely carved out already in 2016 and early 2017. A model without debt effects will have larger problems predicting the house price in case of episodes such as the Financial Crisis. A potential future extension of the thesis could examine the population effects on the house price in a longer horizon, for instance with annual data. Additionally, the effect of the new mortgage regulations on the households' debt would have been interesting to follow in a longer perspective.

An interesting finding was that backdated growth in house prices is an important driver of house prices. Different from earlier studies, the analysis found that both the growth in real house prices one year ago and one quarter ago are important drivers of the current house price. The backdated growth implies that people observe the historic house price when they shape expectations to the future house price. This mechanism can cause a self-reinforcing house price spiral, which can lead to severe house price bubbles. Especially the effect of the growth in the preceding quarter is worrying, as it implies that people react faster to changes in house price growth, than what was found in the earlier studies. In turn, this can in part help to explain why the Norwegian house prices have been increasing so intensely the last decade compared to the OECD average.

According to the preferred ARDL model, the real house prices in Norway were above their fundamental value from mid-2016 until the end of the sample in 2017. With expectations of a rising interest rate, the interest burden of many households will be substantial, which can reduce their consumption dramatically. Thus, the house price development will have a strong impact on the Norwegian economy in the years to come.

References

Anundsen, A. K., & Jansen, E. S. (2013). Self-reinforcing effects between housing prices and credi: an extended version. *Discussion papers No. 756, Statistics Norway, Research Department.*

Barnes, R. (2009). *The fuel that fed the subprime meltdown.* Retrieved from Investopedia: https://www.investopedia.com/articles/07/subprime-overview.asp

Barot, B., & Yang, Z. (2002). House prices and housing investment in Sweden and the UK: Econometric analysis for the period 1970-1998. *Review of Urban and Regional Development Studies*, 189-216.

Birkelund, H. R. (2017, May 29). *Nå er flertallet av norske arbeidstakere uorganisert.* Retrieved from frifagbevegelse.no : https://frifagbevegelse.no/nyheter/na-er-flertallet-av-norske-arbeidstakere-uorganisert-6.158.468089.cdbf5ed3a1

Bjørnestad, S. (2017, February 20). Norske lønnstagere hadde i 2016 lavest lønnsvekst siden krigen. *Aftenposten.*

Bjørnland, H. C., & Thorsrud, L. A. (2015). *Applied time series for macroeconomics.* Gyldendal.

Boug, P., Johansen, P., & Naug, B. (2002). MODAG - en makroøkonomisk modell for norsk økonomi. *Sosiale og økonomiske studier, SSB.*

Brooks, C., & Tsolacos, S. (2010). *Real estate modelling and forecasting.* Cambridge.

Brown, R. L., Durbin, J., & Evans, J. (1975). Technieques for testing the constancy of regression relationships over time. *Journal of Royal Statistical Society*, 149-163.

Christensen, J. (2018, April 6). *Skattereformene i 1987 og 1992.* Retrieved from Store Norske Leksikon: https://snl.no/Skattereformene_i_1987_og_1992

Cochrane, J. H. (1990, April). *A critique of the application of unit root tests.* Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.308.7613&rep=rep1&type=pdf

Corder, M., & Roberts, N. (2008). Understanding dwellings investments.

Dagens Næringsliv. (2017, October 4). *IMF-økonomer advarer: husholdningenes økende gjeld kan gi bakrus.* Retrieved from Dagens Næringsliv: https://www.dn.no/nyheter/2017/10/04/0926/Makrookonomi/imf-okonomer-advarer-husholdningenes-okende-gjeld-kan-gi-bakrus

Davidson, R., & MacKinnon, J. (2003). *Econometric theory and methods*. Oxford University Press.

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and Inference in Econometrics*. Oxford Unisersity Press.

E24. (2018, January 5). Oslo svakeste i 2017 med 6,3 prosent nedgang. Retrieved from https://e24.no/privat/bolig/oslo-svakeste-i-2017-med-6-2-prosent-nedgang/24223165

Eiendom Norge. (2018). *Boligprisstatistikken*. Retrieved from http://eiendomnorge.no/boligprisstatistikken/#nokkeltall

Eitrheim, Ø. (1993). En dynamisk modell for boligprisene i RIMINI. *Penger og Kreditt nr.4/1993*, 288-297.

Engle, R. F., & Granger, C. W. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. In *Econometrica, Vol. 55, No. 2* (pp. 251-276).

Finansdepartementet. (1992, August 31). *Norges offentlige utredninger - Bankkrisen*. Retrieved from https://www.nb.no/items/URN:NBN:no-nb_digibok_2007112804031

Finanstilsynet. (2010, March 3). *Rundskriv - Retningslinjer for forsvarlig utlånspraksis for lån til boligformål*. Finanstilsynet.

Finanstilsynet. (2018, February 28). Utkast til høringsnotat.

Giles, D. (2013, June 19). *ARDL Models - Part 2 - Bounds Tests*. Retrieved from Econometric Beat: Dave Giles' Blog: http://davegiles.blogspot.com/2013/06/ardl-models-part-ii-bounds-tests.html

Gustavsson, M. (2012, June 14). *Hvorfor ble krisen så mild i Norge?* Retrieved from minervanett.no: https://www.minervanett.no/hvorfor-ble-krisen-sa-mild-i-norge/

Hendry, D. F. (1984). Econometric Modelling of House prices in the United Kingdom. *Econometrics and Quantitive economics*, Chapter 8.

IMF. (2003). When bubbles burst. *World Economic Outlook April 2003*.

IMF. (2004). "*The global house price boom*". Washington: World Economic Outlook September 2004.

Iversen, K. O. (2016, February). *En toroms til bare 50.000 kr?!* Retrieved from DNB Eiendom: http://www.dnbeiendom.no/altombolig/kjop-og-salg/boligpriser1/boligpriser-da-og-na/

Iversen, K. O. (2018, January 15). *Hva lønner seg: å leie eller eie bolig*. Retrieved from DNB Eiendom: http://www.dnbeiendom.no/altombolig/kjop-og-salg/boligokonomi/utleie-leie/lonner-det-seg-a-leie-eller-eie-bolig/

Jacobsen, D. H., & Naug, B. (2004). *What drives house prices?* Norges Bank.

Jacobsen, D. H., & Naug, B. (2004b). What influences the growth of household debt? *Economic Bulletin 04 Q3*.

Jansen, E. (2009). Kan formueseffekter forklare utviklingen i privat konsum. *Samfunnsøkonomen*.

Johannessen, R. (2004, February 23). *Husleier i konsumprisindeksen*. (SSB) Retrieved June 2018, from SSB.no: https://www.ssb.no/priser-og-prisindekser/artikler-og-publikasjoner/husleier-i-konsumprisindeksen

Krogh, T. S. (2010). *Credit Regulations in Norway, 1970-2008. Report 37/2010*. Statistics Norway.

Macrotrends. (2018). *The crude oil price - 70 year historical chart*. Retrieved from http://www.macrotrends.net/1369/crude-oil-price-history-chart

Marschhäuser, S. H. (2015, October 15). *Seks grunner til at du bør eie bolig*. Retrieved from Aftenposten: https://www.aftenposten.no/bolig/Seks-grunner-til-at-du-bor-eie-bolig-7138b.html

NHO. (2018, March 9). *Lønnsoppgjøret 2018*. Retrieved from nho.no: https://www.nho.no/Lonn-og-tariff/Lonnsoppgjoret/dette-er-lonnsoppgjoret-2018/?id=88553

Norges Bank. (2015). *Høring - krav til utlån med pant i bolig*.

Norges Bank. (2018, March 15). *Pengepolitisk rapport*. Retrieved from https://www.norges-bank.no/Publisert/Publikasjoner/Pengepolitisk-rapport-med-vurdering-av-finansiell-stabilitet/2018/118/

OECD. (2018). *Analytical house price indicators*. Retrieved from stats.oecd.org: https://stats.oecd.org/Index.aspx?DataSetCode=HOUSE_PRICES#

Perlberg, S. (2013, December 2). *Business Insider*. Retrieved from http://www.businessinsider.com/norways-housing-market-is-a-bubble-2013-12?r=US&IR=T&IR=T

Pesaran, M., Shin, Y., & Smith, R. J. (1999). Bounds testing approaches to the analysis of level relationships.

Pesaran, M., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 289-326.

Phillips, P. C., & Perron, P. (1988). "Testing for a Unit Root in Time Series Regression". *Biometrika*.

Piazzesi, M., & Schneider, M. (2016, July). Housing and Macroeconomics. *NBER*.

Pihl, C. (2017, August 18). *Ny statistikk: Snittalder på førstegangskjøpere er 28 år*. Retrieved from http://www.nef.no/nyheter/snittalder-pa-forstegangskjopere-er-280-ar/

Regjeringen. (2013, December 27). *Kåre Willoch*. Retrieved from https://www.regjeringen.no/no/om-regjeringa/tidligere-regjeringer-og-historie/historiske-artikler/embeter/statsminister-1814-/kare-isaachsen-willoch/id463461/

Skatteetaten. (2018). Retrieved from Formuesverdi av egen bolig: https://www.skatteetaten.no/person/skatt/hjelp-til-riktig-skatt/bolig-og-eiendeler/bolig-eiendom-tomt/formuesverdi/egen-bolig-primarbolig/

SSB. (2010). Fullførte bygg, etter brukerens næring. Retrieved from: https://www.ssb.no/a/histstat/tabeller/17-4.html

SSB. (2016, Juni). *Levekårsundersøkelsen*. Retrieved from Statistisk Sentralbyrå: https://www.ssb.no/bygg-bolig-og-eiendom/statistikker/bo/hvert-3-aar

SSB. (2018, May 28). *Boforhold, registerbasert*. Retrieved from ssb.no: https://www.ssb.no/bygg-bolig-og-eiendom/statistikker/boforhold/aar

SSB. (2018, Februar 1). *Den lave lønnsveksten fortsetter*. Retrieved from ssb.no : https://www.ssb.no/arbeid-og-lonn/artikler-og-publikasjoner/den-lave-lonnsveksten-fortsetter

Woolridge, J. M. (2013). *Introductory econometrics*. South-Western Cengage Learning.

World Bank. (2017). *World GDP growth*. Retrieved from worldbank.org: https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG

Appendix

Appendix A – Definitions and collection of the data material

Variable	Source
The house price index	Data received from Bjørn Naug over email. The data material has been collected from finn.no, NEF, Eiendom Norge and SSB. Extended data was collected from the house price index that is published monthly by eiendomnorge.no. The nominal index uses 2002 as the base year. The average over three months was used to create quarterly values. The data has been deflated with CPI-ATE.
The households' debt	Monthly data published in table 06718 at ssb.no. Extended data received over email from Torbjørn Cock Rønning at SSB. The data has been deflated with CPI-ATE
The interest rate	Data received from Bjørn Naug over email. Extended with data from table 10648 at ssb.no. The real interest rate was calculated through the Fisher equation: $$R = \frac{1+i}{1+\pi} - 1$$ where R represents the real interest rate, i represents the nominal interest rate, and π represents the inflation. Following Jacobsen and Naug (2004), CPI-ATE have been used as a measure of inflation.
The households' disposable income (the wage income).	Quarterly data received over email from Jørgen Ouren in SSB. Total wage income in all sectors (YWW). Deflated with CPI-ATE. In 2015 there was a modification to how Statistics Norway (SSB) measured the aggregated wage income. As a result there was a much more extreme seasonal variance in the data from 2015. To make the series more consistent I smoothed

	the last three years of data with a moving average of four periods.
The housing stock in fixed 2015 prices	Data received over email from Bjørn Naug. Additional data received over email from Pia Tønjum at SSB.
Initiated square meters of housing	Monthly data published in table 10996 (1983M1-1999M12) and 05808 (2000M1-2018M5) at ssb.no. To accumulate the series in levels, the perpetual inventory method has been used: $K_{t+1} = (1 - d)K_t + I_t$ where K represents existing housing stock, d represents the capital depreciation (which is assumed to be constant at 2 percent like in the national accounts) and I represents new investments.
The unemployment rate	Data received from Bjørn Naug over email. Extended with data from *Hovedtall om arbeidsmarkedet 2017* at nav.no. Monthly data has been transformed to quarterly data through a 3 month average.
The households' expectations	Data received from Bjørn Naug over email. Extended data received from Ann Håkonsen at Finans Norge over email.
The tax rate	Data received from Bjørn Naug over email. Extended data collected from "Skattesatser 2018" at regjeringen.no
Core inflation CPI-ATE	Data received over email from Bjørn Naug. The extended data was collected from table 05327 at ssb.no. Mothly data has been transformed to quarterly through 3 month averages.
Total population	Quarterly data from table 01222 at SSB.no
Amount of people in a household	Annual data from table 09747 at SSB.no
Relocations	Annual data from table 09498 at SSB.no
Share of population in central areas	Annual data from table 05212 at SSB.no
House rents from the CPI	Paid rents. Monthly data from table 03013 at SSB.no
House rents – rental market survey	Annual data table 09895 at SSB.no.

	Monthly data from 2006-2012 from table 06257 at SSB.no
GDP	Table 09190 at SSB.no

Appendix B - Overview of the unit root tests

Values that are below the 5% critical value have been **highlighted** for the variables in levels. For the differenced variables the test statistics that are <u>not</u> below the 5% critical value have been **highlighted**.

The unit root tests of the variables

Variable	ADF	5% critical value	PP	5% critical value
houseprice (t)	-3.117 (5)	-3.447	-1.868 (5)	-3.446
income (t)	-3.392 (6)	-3.447	-2.082 (6)	-3.446
INTEREST[24]	-1.565 (2)	-2.888	-1.418 (2)	-2.888
debt (t)	-2.368 (6)	-3.447	-1.555 (6)	-3.446
housestock (t)	-1.805 (6)	-3.447	-0.287 (6)	-3.446
initiatedsqm (t)	**-3.745 (3)**	-3.447	**-13.679 (3)**	-3.446
unemployment	**-3.304 (6)**	-2.888	-2.363 (6)	-2.888
houserent (t)	-2.737 (6)	-3.447	-2.655 (6)	-3.446
expectations	**-3.366 (1)**	-2.888	**-3.370 (1)**	-2.888
Δhouseprice	-3.793 (4)	-2.889	-3.793 (4)	-2.888
Δincome	-3.017 (6)	-2.889	-11.693 (4)	-2.888
ΔINTEREST	-7.015 (1)	-2.888	-7.990 (1)	-2.888
Δdebt	-2.997 (5)	-2.888	-6.227 (5)	-2.888
Δhousestock	**-2.680 (4)**	-2.888	**-1.975 (4)**	-2.888
Δinitiatedsqm	-7.578 (1)	-2.888	-5.144 (1)	-2.888
Δunemployment	-4.326 (4)	-2.888	-16.647 (4)	-2.888
Δhouserent	-3.795 (5)	-2.888	-12.218 (5)	-2.888

---

[24] If a trend is included in the ADF test, non-stationarity is clearly rejected.

Δexpectations	-11.325 (0)	-2.888	-11.359 (0)	-2.888

(t) = trend included in the tests. Lags are in parentheses behind the test statistics. Δ indicates that the variable is differenced once.

Appendix C – Adjusting the expectations variable

Following Jacobsen and Naug's adjustment procedure, EXPEC is calculated as:

ΔE_t	EXPEC 1992-2017
$\Delta INTEREST(1-\tau)_t$	-8.51***
	(-4.65)
$\Delta unemployment_t$	-0.42***
	(-2.97)
E_{t-1}	-0.18***
	(-3.29)
$INTEREST(1-\tau)_{t-1}$	-0.62
	(-1.28)
$unemployment_{t-1}$	0.039
	(1.45)
Q1	0.19***
	(5.48)
Q2	0.07***
	(3.53)
Q3	0.17***
	(5.56)
Constant	0.075
	(0.70)
Observations	102
R^2	0.493

t statistics in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Then the residuals from the regression, e_t is saved. Jacobsen and Naug further modifies the residuals: $EXPEC = e_t + 100 * e_t^3$. They do not provide any explanation why they choose to add the latter term in the expression. From the figure below, the residuals and the EXPEC variable are plotted. The EXPEC variable has much more extreme effects around the peaks. Thus, it is almost working as a dummy for the Financial Crisis and the oil price shock in 2014.
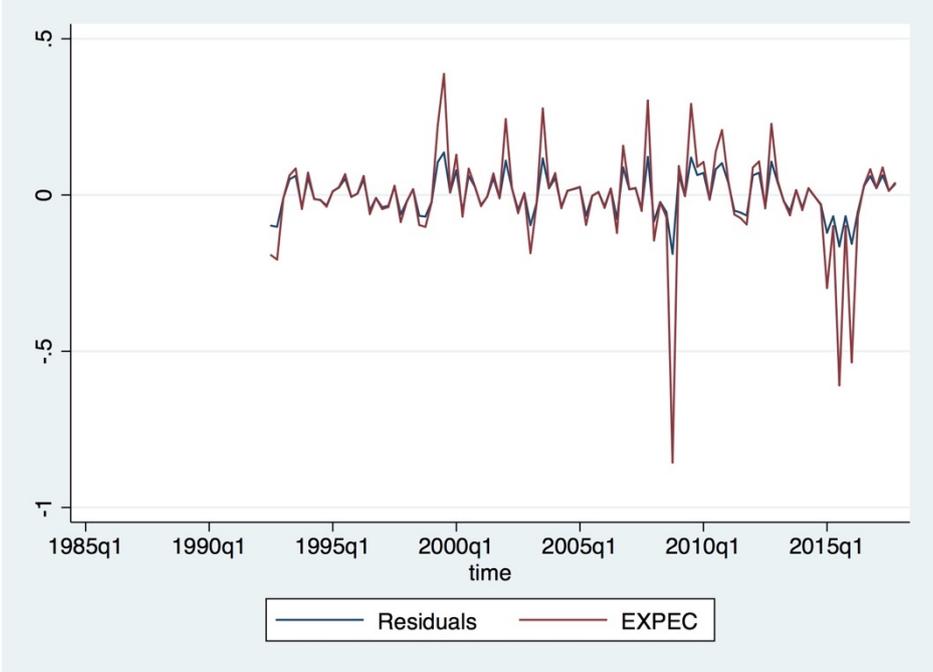
Appendix D.1 – Johansen's tests of the variables in the ARDL

Johansen's test of the variables in model 1:

Rank	Trace statistic	5% critical value	Probability
None	43.27	40.17	0.02
1	21.15	24.28	0.11
2	2.11	12.32	0.94
3	0.23	4.13	0.69

With 5 lags, no constant and no trend. Seasonal dummies and EXPEC as restricted variables.

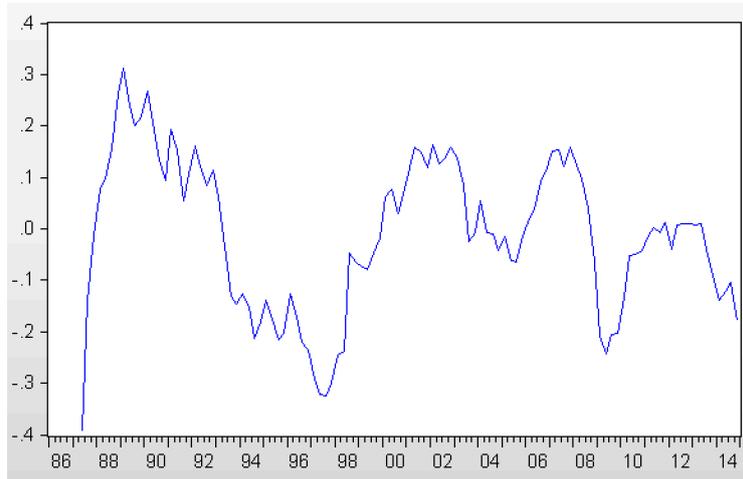Johansen's test of the variables in model 2:

Rank	Trace statistic	5% critical value	Probability
None	79.39	60.06	0.00
1	36.23	40.17	0.12
2	17.35	24.28	0.29
3	2.99	12.32	0.85

4	0.00007	4.12	0.99

With 5 lags, no constant and no trend. Seasonal dummies and EXPEC as restricted variables.
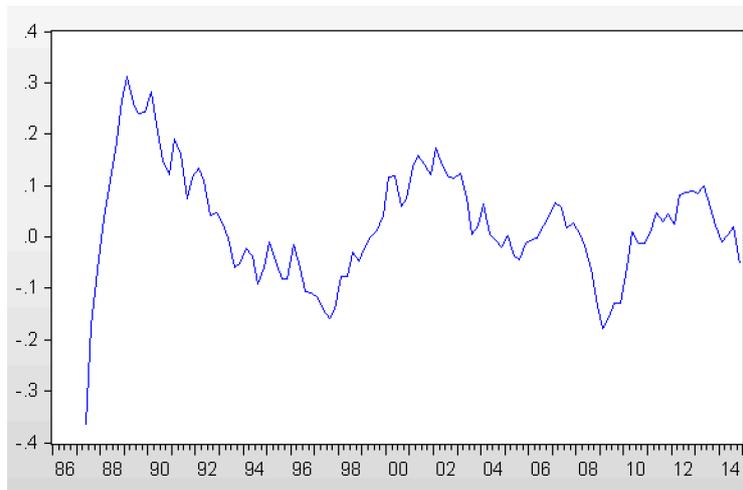
Appendix D.2 – The cointegrating relationships in the ARDL

Cointegration in model 1 – without debt

An ADF test with 4 lags on the cointegration finds that it is highly stationary.

Cointegration in model 2 – with debt

An ADF test with 4 lags on the cointegration finds that it is highly stationary.