NORGES HANDELSHØYSKOLE

Bergen, Autumn semester 2012

LONG-TERM EFFECTS IN THE NORWEGIAN HOUSING MARKET

- an empirical analysis

Harald Haltvik and Stian Øya Bendiksen

Veileder: Zuzana Laffersova

Master thesis in Financial Economics

NORGES HANDELSHØYSKOLE

This thesis was written as a part of the Master of Science in Economics and Business Administration program - Major in Financial Economic. Neither the institution, nor the advisor is responsible for the theories and methods used, or the results and conclusions drawn, through the approval of this thesis.

Abstract

In this thesis we have analysed the Norwegian housing market. On the basis of our examination of possible factors affecting the housing market, we have estimated a model to explain the Norwegian housing prices. Especially, we wanted to test if the state of the economy, categorized by the different phases in a business cycle, had a significant effect on the housing prices We have estimated a short-term dynamic and a long-term solution, where we show the different factors' affect on housing prices.

We start by presenting the development in the Norwegian economy, and the peculiarities of the housing market. To create a basic understanding of housing models and previous research, we present the two main models for housing prices conducted in Norway the last 15 years. We then presented the variables we believed to affect the housing prices and the statistical methods used. The final model was estimated and tested, and proved to be good at both explaining and predicting the housing prices in the period from 1986 to 2011.

In our final model, we could not find any significant effects from the different states in the business cycle. We found that the short-term dynamic in the housing market are affected by the change in housing prices in earlier periods, and the long-term solution is affected by the development in households' real disposable income, housing stock and real interest rate. We also tested the models ability to predict housing prices. In general the predictions were fairly decent, which indicates that the variables have a consistent economic grounding throughout the period. There will although be problems with using the model to predict future housing prices, since the development in the variables are uncertain and often subject of revising in retrospect.

Preface

After four years together at NHH, we have discussed a lot of potential subjects for our master thesis. There was no lack of interesting economic thesis and the choice was difficult. After some discussion, we found out that we shared a common interest for the housing market. We are both in the process of getting involved in the housing market, one is selling and the other one is buying for the first time. With the exceptional development in the housing market in Norway the last decades, it became pretty clear what focus we wanted in our master thesis. We wanted to find the main determinants in the housing market.

During the process of writing our thesis we have been in contact with a lot of helpful and inspiring people. First of all, we want to thank our supervisor, Zuzana Laffersova, for great help and inspiration. She has been at great service, also by helping us getting in touch with other people. We will therefore also thank Svein Olav Krakstad, Eilev Jansen and Zuzanas husband, for their help and guidance. We will also thank SSB for helping us with the gathering of data.

We have learned much about the Norwegian housing market and econometrics from working with this thesis. Neither of us had the sufficient knowledge of the econometric methods nor the software used in this thesis when starting the process. Therefore we have used a great deal of time acquiring this knowledge. By focusing on empirical analysis, we feel that we have applied our knowledge from both finance and economics. The thesis has been both challenging and rewarding to work with it.

Bergen, December 2012

Stian Øya Bendiksen

Harald Haltvik

Contents

Table of figures

1. Introduction

The growth in Norwegian housing prices has been tremendous for the last 20 years. The development in housing prices affects many Norwegians, especially young people entering the market. High prices, equity regulations and stricter lending policies, have made it very difficult for first time buyers. Since 80% of all Norwegian households live in dwellings that they own, the housing prices are essential in terms of personal wealth and personal economy.

The Norwegian housing prices decreased as a result of the financial crisis in 2008, along with most countries in the world. The Norwegian market's quick rebound and pro-longed incline in the post-crisis period is quite a unique case. Economists and experts are debating whether or not the increased prices is caused by changes in fundamental values or a bubble taking shape in the housing market.

In this thesis we will focus on the determinants in the housing market in the period from 1986 to 2011. Presented determinants will be based on economic theory, previous research and our own statistical analysis. First we will look closer at the economic development in Norway and the Norwegian housing market in the given period, while creating expectations of possible determinants. We will categorize every quarter as one of the four states in a business cycle. Especially, we want to test if the state of the economy, categorized by the different phases in a business cycle, had a significant effect on the housing prices. As far as we know, this has not been conducted with Norwegian data before.

We will then construct an Error Correction Model explaining the Norwegian housing market with both short-term and long-term effects. We will use the statistical software STATA in all of our econometric estimations.

The major limitations in our thesis have been the short period of time for writing the thesis, and the availability of data. We needed to do several changes due to missing data, and the data gathering process was more demanding than we had anticipated. To start with, our knowledge of advanced econometrics and use of statistical software was rather limited. This resulted in much time spent acquiring the required knowledge to conduct the analysis.

In the thesis, we start with presenting the development in the Norwegian economy in chapter 2 and business cycle theory in chapter 3. We then present the peculiarities of the housing market in chapter 4, before we look closer at previous research and housing models in chapter 5. We then start evaluating the possible factors affecting the Norwegian housing market in chapter 6, and present the statistical methods we have used in chapter 7. In chapter 8 we start building our housing prices model, and then we estimate the final model in chapter 9. In chapter 10, we comment on the validity of our model in chapter 10, before we make our conclusions in chapter 11.

2. The Norwegian economy

2.1 The period 1990-2012 – a short summary

When entering the 1990s, the Norwegian economy was in the strongest recession since the Second World War. The Norwegian recession in the late 1980's was accompanied by an international economic upturn and strong growth in the traditional exports, which indicates that the recession can be related to domestic conditions. The previous period with increasing debt level among households, the substantial increase in the real interest rate, increasing unemployment and decreasing housing prices resulted in reduced household consumption. The lending boom was followed by big losses and crisis in the banking industry (Benedictow, 2006). An international recession prolonged the already strong Norwegian recession to the end of 1992.

Earlier, from 1978 until 1990, the Norwegian exchange rate was held fixed to a currency basket, which represented the composition of Norwegian foreign trade. In 1990 the krone was tied up to ECU, the precursor to the euro. With increased movement of capital during the 1990s, it became harder to keep the exchange rate fixed. A small difference in the interest rate would result in great movement of capital and an increased pressure on the Norwegian currency. Since the Norwegian economy normally developed desynchronized with the foreign countries, the monetary policy would often have a procyclical effect (Benedictow, 2006). This implies that the monetary policy enhances economic upturns and downturns, instead of having a dampening effect. Therefore, the Norwegian economy experienced a conflict of interest between the monetary policy and the currency target with a fixed exchange rate.

After the turmoil in the financial markets in 1992, many European countries were forced to depreciate their currency and investors expected Norway to follow. This resulted in capital flight from the krone and the Norwegian central bank was forced to raise the key policy rate to make it more attractive to invest in the krone, even though the Norwegian economy was in a downturn. This made Norway leave their fixed rate, but kept the goal of obtaining stable exchange rates with the European currencies.

By the start of 1993, the Norwegian economy entered a long lasting upturn, supported by reduced interest rates, increased public consumption and an international economic turnaround. The household economy was strengthened after years with mortgage down payments (Eika, 2007). The years of recession in Norway contributed to a lower inflation than the nation's trading partners. Norway's increased competitiveness boosted the internationally exposed industries. Along with growth in earnings and increased capacity utilization, the employment increased by 230 000 jobs in the period from 1993 to 1998. The unemployment rate was halved from 5% to 2,5%. (Benedictow, 2006). The growth in Norwegian economy declined in 1998 as a result of the turmoil in the international capital markets related to the so-called Asian financial crisis in 1997. The strong decline was caused by a fall in oil prices and the doubled interest rate, but the downturn did not affect Norway's most important trading partners to large extent, and the upturn continued in the US and Europe.

With the increasing revenues from the oil industry, it became harder to restrict the government spending. The increased economic activity created expectations of a future raise in the interest rate, leading to an appreciated krone. Towards the end of the 1990s, the monetary policy was aiming for a low and stable inflation. By targeting the inflation through the monetary policy, the government believed that this change of policy would better suit the level of activity in the Norwegian economy.

In March 2001, the inflation targeting was formally introduced along with the "budgetary rule" concerning the use of revenues from the oil industry. The rule states that the revenues should be phased gradually into the economy. By only using the annually expected return (4%) of the "Government Pension Fund", where the revenues are deposited, future generations could also benefit from the fund. The rule is flexible, which means that the government spending can be adjusted depending on the current situation in the economy. The Norwegian central bank got the main responsibility of targeting the inflation at 2,5%. An inflation target results in a countercyclical monetary policy, and aims to reduce fluctuations in the economy.

At the end of the 1990s, a stock bubble was building in the Norwegian and international stock markets. The bubble was caused by high expectations of the return on investments in the information and communication technologies. Eventually it became clear that these expectations were not real and investors wanted to sell their shares. This resulted in a strong decline in prices and other sectors in the economy were affected as well. The American economy went into a recession in 2001 and the rest of the OECD countries followed.

After a long period with economic growth, a pressure was building in parts of the Norwegian economy in 2000, resulting in a higher income growth in Norway compared to the country's trading partners (Benedictow, 2006). Analysis performed by the Norwegian central bank in spring 2002 revealed an increasing risk of inflation in the Norwegian economy. At the same time there was a recession and declining interest rates internationally. A relatively high Norwegian interest rate resulted in an appreciated krone, and by 2003 the krone had appreciated 20% since 2000. The reduced competitiveness along with the recession abroad, made the Norwegian industry suffer and many workers lost their jobs. The Norwegian economy entered a recession in 2002. The unemployment rate had increased continuously for 4 years and reached a peak at 4,7% in 2003(Eika, 2007). The strong krone, low inflation abroad, high oil prices and the increased import from low-cost countries such as China, contributed to a very low inflation in Norway. As a result, the Norwegian central bank lowered the key policy rate from 5,25% in December 2002 to 1,75% in March 2004. This stimulated household consumption and the internationally exposed industries improved their cost competitiveness through a depreciated krone. The strong decline in interest rates made the recession short and moderate and the economy reached the cyclical trough in early 2003.

The following economic upturn in Norway and internationally, together with the weaker krone, turned around the negative development in the industry during 2004. The previous recession indicated that the Norwegian interest rates could not deviate significantly from the interest rates abroad, without making an impact on the currency and the industry. In retrospect, the Norwegian interest rates have only deviated to a small extent from the euro-countries (Benedictow, 2006).

The annual investments in the oil industry increased significantly in 2003, and in the period from 2002 to 2006 the annually investments increased by 60% (Eika, 2007). The oil price doubled from 2003 to 2006 and the oil fund tripled in value from 2003 to the end of 2006. At the same time, the world economy improved and traditional Norwegian exports picked up. The growth was especially high in China and India, and their demand after industrial commodities increased. This benefitted the Norwegian export of industrial commodities. The trade-off gain from cheap imports and increased exports laid the foundation for the significant increase in real wages. In the period from 2003 to 2006, the real wages increased by a total of 9%. Because the turnaround in 2003 and the following growth were driven by a number of factors, it had an impact in most sectors in the Norwegian economy. From 2003 to 2007, there was an increase in household consumption by 22%, an annual growth of 5,5%, and investments in fixed capital increased by 14% per year. In the same period, the employment increased by 215 000 people, but the number of unemployed only fell by 51 000 (SSB, 2012:3). The change in the work force's age composition along with immigration played a major role.

The key policy rate remained at a low level, but the Norwegian central bank started raising it gradually from the summer of 2005. The economic upturn concerned the bank because they expected the inflation to increase over time. It became clearer that the capacity utilization was very high, and as a result the key policy rate was raised more frequently from the fall of 2006. In September 2007 the key policy rate had raised to 5% (Eika, 2007).

Alarming numbers from the American housing market was reported in 2006 and 2007. The default on sub prime loans dragged the financial sector into a crisis and asset prices decreased sharply, affecting the entire economy. When Lehman Brothers declared bankruptcy in august 2008, the financial crisis was a fact. Norway reacted with an expansionary fiscal policy and established financial rescue packages for the banking industry.

During the financial crisis, both the interest rate and currency development contributed with an expansionary effect in the Norwegian economy. The Norwegian central bank's low key policy rate was however offset by the financial turmoil in the international markets. The turmoil resulted in abnormally high premiums in the money market interest rate (Ministry of Labour, 2012). The premiums were especially high at the end of 2008, but it also increased through the fall of 2011. Household consumption growth was considerably lower in 2008 and 2009 compared to the previous years. The turmoil in the international markets has probably had a dampening effect on consumption and has kept the savings rate relatively high. The consumption level also affected imports, which showed low growth compared to pre-crisis figures (Ministry of Labour, 2012). There was no growth in fixed capital investments in 2008, and in 2009 and 2010 the investment level decreased. In 2011, the growth was positive, due to the strong house market and considerable investments in the energy industry.

The import-weighted exchange rate experienced a strong depreciation during the fall of 2008, but the majority of the depreciation was reversed during 2009. For the last years, the krone has appreciated and it is viewed as a secure investment due to the financially solid and stable development of the Norwegian economy. Norway was mildly affected by the financial crisis compared to other countries, but the GDP decreased towards the end of 2008. The financial crisis caused GDP mainland to fall for four consecutive quarters, and the annual growth in GDP mainland turned negative for the first time since 1988. The total decrease in volume these four quarters was about 3%[1]. The decline flattened during the summer of 2009, and since then the mainland GDP has shown an annual growth of 2,5% (Ministry of Labour, 2012). The GDP level before the financial crisis was obtained at the end of 2010, but there was still a big difference in development among industries. If we deduct the gross product of government investments from GDP, the pre-crisis level was not obtained until summer 2011. The household's real income growth has been stable around 3,5-4% in the period 2008-2012.

The Norwegian exports are still struggling after the international downturn and the industry is still experiencing negative growth. Some industries such as fishery products contribute positively to the overall export. Despite the active expansionary fiscal policy, the number of employed fell by a total of 16 000 in 2009 and 2010 (SSB, 2012:3). At the end of 2010 the situation improved and the employment increased and the growth continued through 2011 (Ministry of Labour, 2012).

2.2 Economic outlook

The economic upturn we are experiencing now, is expected to continue for at least four more years (SSB, 2012:2). High growth in demand from the oil industry together with low interest rates are the main forces in the market development. The activity growth in the economy is however likely to be more moderate than earlier upturns.

Weak growth prospects internationally affect the development in Norway negatively. The international economy is still struggling after the financial crisis and the fiscal crisis many countries are still experiencing. It is expected to be less private investments in the mainland

[1] The GDP is seasonally adjusted basic values measured as fixed prices.

economy during this upturn than in previous upturns. Weakened Norwegian cost competitiveness and low international growth will result in modest growth in Norwegian exports, except for petroleum. The development is expected to improve with time, and the pre-crisis export level is predicted to be obtained in 2015 (SSB, 2012:2).

The strong krone has contributed to a low interest rate level, and it is expected to remain at this level towards the summer of 2013. The money market interest rate is expected to reach 4,3% at the end of 2015, which implies mortgage rates over 5%.

Low profitability in the competitive industry is assumed to entail a lower income growth the following year, compared to the current 4,2%. When the growth eventually accelerates internationally, it will result in a higher income growth and a higher inflation. The growth in real income will probably remain at a high level the years ahead. High growth in real wages, employment and social benefits has contributed to an increase in the households' real disposable income. The mortgage interest rates are expected to decrease next year due to low international interest rate level, a strong krone and low inflation. This will probably result in growth in both income and consumption. The key policy rate is expected to increase gradually from late 2013. An increasing interest rate is assumed to have a dampening effect on the growth. The predictions is based on the assumption that the uncertainty regarding the development in both the Norwegian –and international economy will be reduced (SSB, 2012:2).

The investment level in the oil industry is expected to remain high the next years, but with a lower growth than last year's level of 13%. SSB (2012:2) has based the prediction for the forthcoming years on an increased growth in public demand and a continued high growth in social benefits. Therefore, we can characterize the fiscal policy as moderate expansionary. The employment has increased by about 100 000 workers from 2010 to 2012 (SSB, 2012:2), and the economic upturn will probably result in continued growth. The unemployment rate is predicted to be 3,1% in 2012.

In this chapter, we have seen that the Norwegian economy has been through a several upturns and downturns the last three decades. We will now take a closer look at the different phases of the business cycles.

3. Business cycles

Burns & Mitchell (1946, referred in Klovland, Helliesen and Kvåle, 2012) states, “the business cycles are a type of fluctuations found in the aggregate economic activity of nations that organize their work mainly in business enterprises.” These business cycles can be seen as either classical cycles, or growth cycles. Classical cycles have turning points when the trend cyclical curve’s derivative is 0. Growth cycles on the other hand, has turning points when the cyclical curve’s derived is the same as the trends derived.

3.1 Different phases and economic indicators

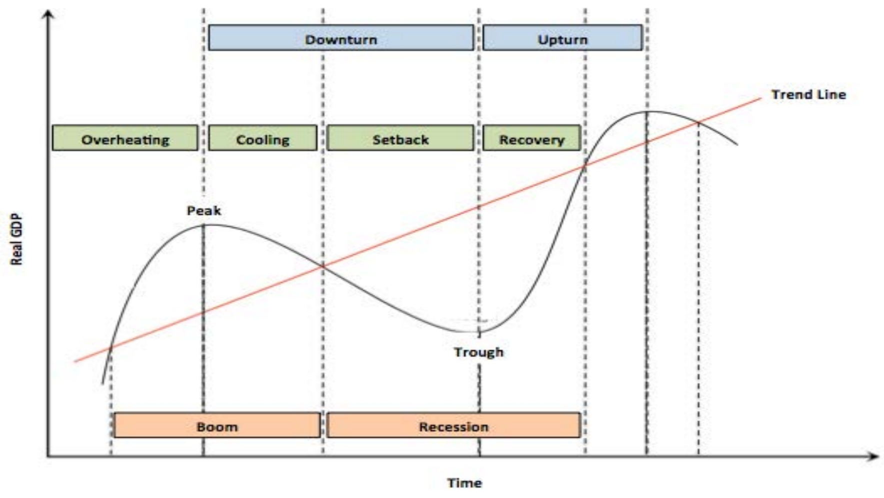
A business cycle can be divided into different stages. There are multiple indicators and methods we can use. In SSB’s description of the business cycles, the GDP mainland in fixed prices is preferred (Eika, 2008). The development in employment is also an indicator that can be easily associated with business cycles. The description of the cycle can be divided into two factors, activity level and development. When the activity level is high, we can identify this as an “economic boom”. This state coincides with a low level of unemployment. The opposite, when the activity level is low, and the unemployment is relatively high, we identify as a recession. The two different states can be associated with the economy being in a good or bad condition. When there is positive growth in the economy, we can identify this as an economic upturn. In this state, the unemployment will normally fall. When the growth is negative, we can identify as an economic downturn and the unemployment is likely to increase.

The different states in the economy can be used to determine whether the economic situation is improving or deteriorating. Based on the GDP mainland we can identify the development by comparing it to the trend. The trend is representing the underlying long-term development in the economy. When the GDP mainland is above or below the trend, the economy is in a boom or recession, respectively. The transition point between an economic downturn and upturn, or the opposite, can be identified as the cyclical trough or peak, respectively. There is a downturn when the business cycle is moving from peak to trough, and an upturn when it moves from trough to peak. The GDP numbers should be seasonally adjusted, so that the classification is not affected by random fluctuations. Based on the above, the economy can at any time be identified in one of four states:

- Economic upturn, but recession – “Recovery”
- Economic upturn, boom – “Overheating”
- Economic downturn, but boom – “Cooling”
- Economic downturn, but recession – “Setback”

The different phases can be summarized in the figure 1. The black line represents the gross domestic product and the red line represents the trend.

Figure 1. Phases in a business cycle



3.2 GDP – not an indicator without problems

GDP is often used as an indicator of the development, wealth and activity level in the economy. When using GDP as a volume indicator, we face methodological problems, since a development in value must be able to be decomposed in volume and price (Eika, 2008).

GDP has several weaknesses as a wealth indicator in the economy. It says nothing about the distribution of income in the society, because GDP is mainly based on market transactions and will not capture the value created at home. This is very important to emphasize in periods when women’s labor force participation is increasing. Also, the value of leisure is not included, so an increase in holiday entitlement, with no change in productivity, will result in a corresponding decrease in GDP.

Furthermore, consumption of natural resources is not considered as a cost. Pollution is a good example. We are not living off GDP, but consumption and good health. GDP can be interpreted as the income generated domestically and as what is available for consumption and saving (consumption in the future).

GDP can also be problematic when used as an activity indicator. In certain industries, the employment can be very low compared to the industry's production value. The development in volume will therefore be of little significance concerning the development in activity. The petroleum industry is an example of such an industry. In this thesis we have solved this problem by focusing on the GDP mainland (GDP which excludes petroleum production and international shipping) when analyzing business cycles.

According to Thorbjørn Eika (2008), the best way to describe the development in GDP is to look at both value and volume. It is the development in volume that determines the labor development. On the other hand, it is the created values we use for consumption purposes. How much we can consume depends on the price ratio between the goods we sell and the goods we buy. The development in volume is the most important. The development in value is not that interesting as long as the prices grow at an equal pace. Another problem we have to keep in mind when working with GDP data is the risk of revision of data at a later point in time.

In this chapter we have taken a closer look at the different phases of the business cycle, which we will use later in the thesis to categorize the last 30 years. We will then use this data to test if it might have any impact on the development in the housing market, which we will look further into in the next chapter.

4. The Norwegian housing market

There has been a substantial development in the Norwegian housing market since the Second World War. We will though emphasize the last three decades, since these are the periods used in the analysis later in the thesis. In this chapter we will also introduce the peculiarities of the housing market, focusing on the Norwegian market.

4.1 Market development

The Norwegian housing market has experienced major alterations and different regulations since the 1940s. From 1940 until 1969, the Norwegian housing market was affected by strict regulations, for example price freezing.

Figure 2, Regulations in the Norwegian housing market

Type of housing	Period	Type of regulation
Rental dwellings	1916-1935	Rent control on some types of flats.
	1940-2010	Rent control on some types of flats.
	1976-1983	Condominium conversion forbidden.
Owner-occupied dwellings	1940-1954	Prize freeze.
	1954-1969	Price regulations.
Housing co-operatives	1940-1954	Price freeze.
	1954-1982/88	Price regulations on new/old flats.
	1976-1983	Condominium conversion forbidden.

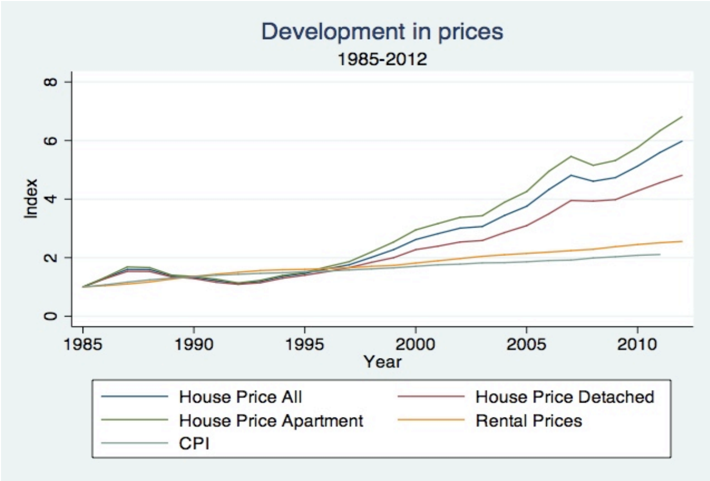
The regulations were so strict that in the period from 1940 to 1954, the nominal prices only increased by 15%, while the consumer price index increased with as much as 90% in the same period. However, this was not the general case in all cities. Prices in Kristiansand increased more than in other cities, and the prices in Oslo were very volatile during the Second World War (the small number of observations made in Oslo during this period may have caused this). When the price freeze regulations were revoked, the prices made a positive jump in the years 1954-55. However, the market was still regulated until 1969, but the regulations were gradually revoked and the nominal prices experienced a significant increase from 1954-1969 (Eitrheim and Erlandsen, 2004).

Norwegian housing prices have had a tremendous growth since 1970. In general, the nominal prices increased by 1300% from 1970 to 2003. The growth in house prices equaled the growth in the consumer price index in the 1970s, leaving the real prices practically unchanged (Eitrheim and Erlandsen, 2004). Real housing prices increased strongly during the liberalization of the credit markets in the 1980s. The decade was characterized by a credit-financed boom. The economy experienced an excessive supply of liquidity from the Norwegian bank, which resulted in a doubling of the money stock (Hodne & Grytten, 2002). The combination of a politically controlled low interest rate and the liberalization of the credit markets caused a rise in demand for consumer goods and dwellings.

This economic boom peaked in 1987 and the Norwegian economy entered a recession lasting until 1993. The recession was reinforced by the banking crisis in the beginning of the 1990s. The recession had a negative effect on the housing market, both in nominal and real terms. The real housing prices decreased by as much as 40% until the market rebounded in late 1992 (Jansen, 2011). Since 1993, the housing prices in Norway have risen unlike any other consuming good. SSB reports that since they started developing their housing price index in 1992, the increase in prices has been 340%. During the same period the consumer – and building price index only increased by 47% and 89%, respectively (SSB, 2012:1). If we base our calculation on numbers from Norwegian Association of Real Estate Agents (NEF), the nominal prices have risen by 434% in total over the same period.

The financial turmoil in the world economy during the crisis in 2008 resulted in decreasing house prices in most countries, but the Norwegian prices have continued to increase, apart from the decrease of 8% in real housing prices in 2008 (Sættem, 2012). The fall in housing prices that many people thought would be a correction on the high price level was already obtained in 2010. In 2010 and 2011 the growth in prices was 8.3% and 8.0% respectively (SSB, 2012:1). This development can largely be explained by income growth, supply of new housing and the banks' interest rates. The sharp increase in prices the last two decades has made many market analysts and economists speculate whether this development is based on structural and fundamental changes or if the growth in prices is due to a bubble taking shape. Although, in the last couple of months there has been indications that the market is slowing down and that the prices are stagnating (Becker, 2012).

Figure 3. Development in nominal prices.1985-2012 (NEF.no)



4.2 Outlooks in the housing market

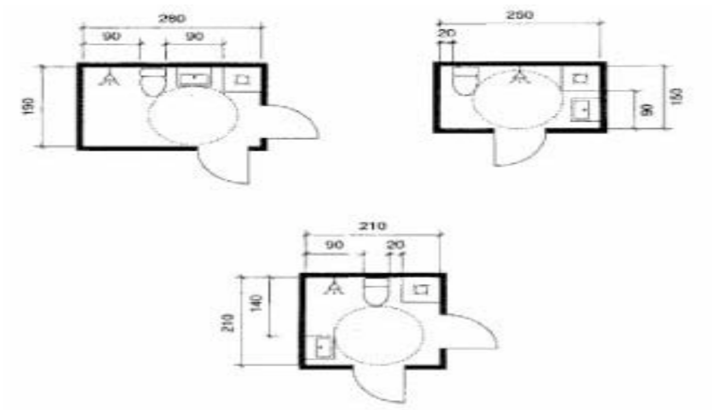
The Norwegian housing prices were 23% higher in the second quarter 2012 than before the financial crisis. According to SSB (2012:4), the last year's growth in prices is mainly due to strong population and income growth along with low interest rates. The increasing prices have contributed to a sharp increase in initiated dwellings. The prospects of low interest rates and high income growth entail a continued and unaffected growth in prices and housing investments. SSB (2012:4), expects that the growth probably will decline when the interest rate level and building costs increases in the future.

4.3 Government regulations and policies in the Norwegian housing market

When building new houses and apartments in Norway, there are several different regulations to consider. To build new dwellings you need available land. We would think that this is the smallest problem in Norway, but in the cities this is a very big issue. Most people want to live in urban areas, but there exists several regulations to secure recreational and agricultural areas. This means that developers have trouble finding suitable land for new profitable projects (De Rosa & Horjen, 2012). One possible solution is upgrading the ways of transportation in and out of the biggest cities, especially the Oslo area. In Norway, each municipality decides which areas are regulated. There are few standard procedures, and each case is treated individually. The long process time slows down the development of new housing areas.

In 2010 there was a change in regulations of technical requirements for new houses and apartments in Norway (Byggteknisk forskrift, 2010) called TEK10. These regulations were made to help facilitating wheelchair users, but requirements regarding inside storage room, free floor space (no furniture) and bathrooms were also included. The regulations are very strict and entail increased costs of planning and building (Byggteknisk forskrift, 2010).

Figure 4. Examples of bathrooms that meets the space requirements for turning a wheelchair (Bolig-abc, 2012).



Eksempler på bad med plass til vaskemaskin. Mål: 1:100.

There is a substantial amount of immigration into the Oslo area and the nearby areas. One problem is that a lot of the households in this area consist of people living alone. This results in an extraordinary need of many small apartments. At the same time, a government policy states that 50% of the apartments being build in the Oslo area needs to be over 80 m^2 (DNB Eiendom Nybygg, 2012). These apartments are very expensive, and many people end up renting a home or living in shared housing. This increases the demand and price of smaller apartments further.

In addition, requirements of better isolation have raised the costs for the developers (Wold, 2009). Together with the requirements mentioned above, this has contributed in raising the costs of building new houses and apartments significantly. These costs have raised the prices of new properties for the consumers correspondingly. Some of the requirements can benefit the consumers in the long run, e.g. better isolation resulting in lower electricity costs, but the problem is that the raised prices of new apartments also will contribute to raise the prices of used apartments. The used dwellings are not faced with the same new requirements, and the price increases without any costs.

4.4 The economic differences between renting and owning in Norway

To illustrate the effects of taxes and regulations in Norway, we will show a numeric example concerning the differences of owning and renting an apartment.

In Bergen, we can buy a nice 3-room apartment for approximately 2,5 million NOK (finn.no, 2012:1). A normal rent for such an apartment is approximately 12 000 NOK per month (finn.no, 2012:2), which would be the same as we would pay if we rented the apartment. To simplify and include a small margin of safety let us say that the effective interest rate is 5%. We simplify things buy assuming 100% debt financing, to avoid the requirement stating that we need 15% in equity to get a house loan in Norway. Let us assume we earn enough to get the full tax deductibles from the interest rate payments. Since the taxable value of our primary residence should not exceed 30% of market value in Norway, while the taxable value of our liabilities are 100% of real values, wealth tax is avoided. Since the apartment in our example is located in Bergen, we do not pay any property tax. Every municipality decides if, and how much, property tax they want to apply. The property tax in Norway must be between 2-7‰, which is 5000-17500 NOK per year for the apartment in the example.

Let us further more assume that there will be 1500 NOK per month in shared costs for the joint ownership and public costs. These are costs that do not occur when we rent instead of owning. For the ease of this example, we will assume that the growth in housing prices will equal the maintenance costs over time even though this has not been the case in the Norwegian housing market the last 25 years (SSB, 2012:3).

This gives us the following calculations to compare owning and buying in an easy way.

Figure 5. Price example owning vs. renting

Rent per month	12 000	
Price of property	2 500 000	
Interest rate	5 %	
Other costs when buying	1 500	
Costs per month	**Buy**	**Rent**
Interest costs	10 417	
Tax reduction	2 917	
Other costs	1 500	
Total costs per month	9 000	12 000

As we can see, we have a benefit on 3000 per month when buying instead of renting. The benefit will depend on the rental price, interest rate and other factors in each individual case.

The Norwegian mentality is very pro owning compared to the neighbour countries and most Norwegians want to own their own house in order to get a real “home-feeling”. This is also shown in the statistics, where about 80% of Norwegian households owned their own home, while about 60% did the same in the rest of Scandinavia in 2009 (Eurostat, 2012). In addition to the tax benefits mentioned above, you can rent out a part of your dwelling that you live in yourself, tax free, as long as the rent is less than half of the dwelling’s full rental value. Any possible surplus occurring when selling your house is also tax-free, as long as you have lived in the house for at least one of the last two years.

To truly compare the costs of renting with the costs of buying, we should include the risk of owning and the opportunity cost for the investment. The need for housing exists no matter which of the two possibilities are chosen. Often, the decision is whether or not to enter the market. Most Norwegians want to own their own house at some point in their life, and the biggest risk is the market timing. When entering the market for a short period of time, we have to mind the significant transaction costs, but in the long run these costs are negligible. The opportunity costs are important, at least the equity part of the investment. Although it is not very likely that you will get the same favourable financing terms when financing alternative investments, as when investing in the housing market.

4.5 How to measure housing prices

There are two main problems with measuring prices in the housing market. First, the housing stock is heterogeneous. There are several factors that are important when determining the value of a property. Even two identical apartments will vary in value according to location, so it is very difficult to find the best way of measuring the general development in the housing market. Housing prices also vary largely due to e.g. size, type of house, standard etc. Second, each property is traded very rarely and it is therefore difficult to follow the individual properties’ development in prices since the last trade. The owners may have done significant improvements increasing the quality and hence the price. The dwelling can also be subject of depreciation if not maintained, causing the price to decrease.

When taking these problems into account, there are three usual ways of constructing indices for housing prices (Klovland, Helliesen and Kvåle, 2012). Using these different methods, we need to be aware of problems connected to each method and its underlying assumptions.

One method tries to follow specific properties through repeated sales, as mentioned above. This is called the "repeat-sales-method". The point is to use the available price information on a specific property at different times. This way you have a pair of transactions including data from the two sales, price p_1 and p_2, at time t_1 and t_2. Indices will then be calculated by using a regression where the change in prices in the transaction pairs are estimated as a function of time dummies. The problem with this method is that it requires a lot of data over a long period of time. It is also difficult to identify any improvements or depreciations in the time between sales.

The second method is making "hedonic housing price indices." You estimate the price as a function of different attributes in relation to the property, like location, standard, etc. Then you calculate the price development for a standard house with constant quality by adjusting for changes in attributes. The problem with this model is to identify and include all attributes that affect the price.

The third method is the "average price" for all house transactions. This is often made as a price per square meter index. The most common housing price index in Norway, "ECON – Eiendomsmeglerbransjens boligprisstatistikk", is made this way. The main index can be divided into different sub indices for different types of houses, locations, etc., but these classifications might be very general. The range of houses sold might also vary over time, and could complicate the construction of the index. The main problem is the lack of consideration of the heterogeneity in the housing market.

When calculating housing price indices, it is also important to notice possible national peculiarities. In Norway, seasonal adjustments are important to see the real development in the housing market in the short-term. As an example amongst other seasonal effects occurring through the year, the housing prices almost always increase from December to January (Bakken, 2010). It is also important to consider inflation in order to see the real development in prices over a long period of time. When looking at the development in the

Norwegian housing market, we need to remember that the index is made as an average price for all house transactions. We also need to know whether or not the index is seasonally adjusted, and if it is specified in real or nominal prices.

4.6 Supply and demand in the housing market – macroeconomic theory

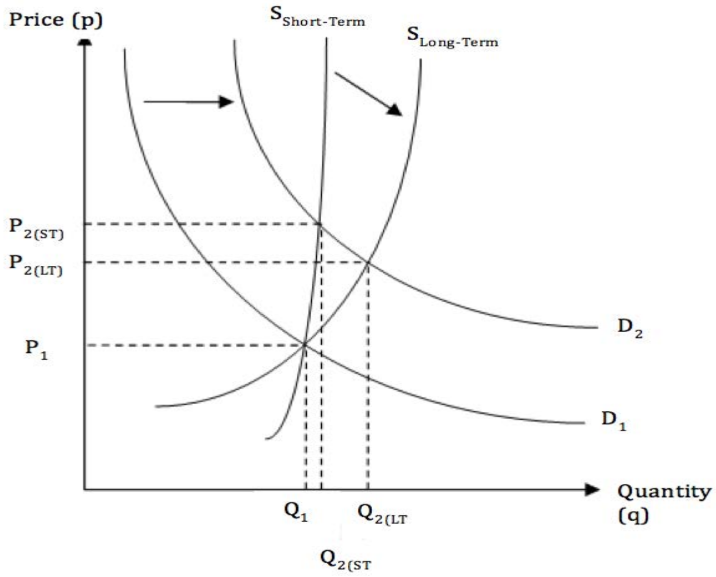
There are different theories and opinions about the future development in prices in the housing market. We will now look further into which factors are affecting the housing prices. The price of any product is determined by supply and demand of the given product. A market for a specific product consists of buyers and sellers of the particular product. The demand curve tells us what quantity of the product the consumers are willing to buy at the given price. The curve is downward sloping, which means that when the price of the product is falling, the quantity demanded by the consumers increase, with the exception of Giffen goods[2]. More customers are the willing to buy the product. The supply curve describes the quantity the suppliers are willing to supply at the given price. The curve is upward sloping, meaning that the suppliers want to produce more when the price increases. The opportunity cost of not producing more of the product is increasing when the price increases.

The Norwegian housing market can be described as less homogeneous, since rather few dwellings are identical. Especially if we compare the housing market with markets for standardized consuming goods. When the market is subject to a boom in demand, the supply curve will react differently in the housing market rather than in a market for ordinary consumer goods. In ordinary markets, manufactures can quickly adjust their supply when unexpected turns in demand occur. This is not the case in the housing market, where the manufacture process is considerably longer than in markets for standardized consumer goods. The supply of houses will therefore be inelastic in short term, as a result of the limited capacity in the construction industry. The newly built houses will represent a relatively low share of the total housing stock, also due to the low capacity. Consequently, it will take time before the total supply of housing is adjusted to the increased demand. The housing market could therefore end up in a situation where the short-term increase in prices

[2] A Giffen good is a good which people consume more of as the price rises. This is caused by the income effect dominating the substitution effect.

can be larger than the long-term increase. A long-term model describing the housing market should therefore include determinants that describe the development in housing stock as well as the development in building costs and land prices (Jacobsen and Naug,, 2004:2). The housing demand consists of two components: the households' demand for housing for living purposes and demand for housing for pure investment purposes. It is reasonable to assume the first component represents the bigger share of the total demand.

Figure 6. Price dynamics in the housing market (Klovland, 2012).



Initially, the market equilibrium is in the intersection (q_1, p_1). The boom in demand causes the demand curve to shift from E_1 to E_2. The sudden increase in demand causes the short-term price to rise from p_1 to p_2. The short-term equilibrium is in the intersection between the new demand curve and the short-term supply curve (q_2 (ST), p_2(ST)), where the housing prices has increased significantly while the housing stock has only increased to a small extent. Due to the limited capacity in the construction industry we can see that the short-term supply curve is steeper than the long-term curve. The increase in prices causes more housing starts as a result of improved profitability. This is illustrated in the figure by a gradual shift from the short-term equilibrium to the long-term equilibrium. The housing stock will continue to increase as long as the housing prices are higher than the long-term equilibrium (q2 (LT), p2 (LT)). The short-term effect on prices is larger than the long-term effect.

We have now presented the housing market, and tried to explain the main peculiarities in the Norwegian market. In the next chapter, we will take a closer look at previous research of the Norwegian housing market and their resulting econometric housing price models.

5. Previous research and econometric housing price models

5.1 MODAG / KVARTS – model

MODAG is a model for the entire economy developed by SSB. This model consists of many sub-models, from which one of them relates to the development in housing prices. Modag is used by the Ministry of Finance in relation to the national accounts and in other cases when estimating and analyzing annual macro economic figures and forecasts, in both short –and long-term aspects. The model has been revised several times through the years. We will focus on the latest version published in late 2008 by Pål Boug and Yngvar Dyvi.

The housing prices in the second hand market in the MODAG-model are modeled from the market demand, while the changes in housing stock are modeled from the market supply. They base the determination of the market price on a given quantity of real capital. The price on second hand housing is the price that clears the market. We will only focus on the housing prices in the second hand market, hence the market demand. The demand of housing (K^E) is determined by the household's disposable real income (Y) and the user price of the housing, which is the cost of holding one unit of housing in one period. The user price is affected by the real interest rate after tax(r) and the housing price (P_K)

The aggregated demand for housing can therefore be expressed as:

$$\boldsymbol{K = K^E(P_K, Y, r)} \qquad \textbf{(1)}$$

Increased housing prices and/or increased real interest rates after tax will result in a decrease in demand for a given Y. An increase in disposable real income will result in a rise in demand for a given level of P_K and r.

In short-term we assume the housing capital as given. Equation (1) can therefore be inverted, expressing the housing price that clears the market. In other words the price that makes the demand for housing equal to the given supply of housing.

$$\boldsymbol{P_K = P_K(K, Y, r)} \qquad \textbf{(2)}$$

For a given amount of housing capital, the price will increase with higher income and decrease with an increase in the interest rate. The price will fall if the housing capital increases. This short-term relationship is also explained graphically in figure 6 in the chapter regarding the price determination.

The long-term equilibrium price is determined in the intersection between the aggregated supply curve and the aggregated demand curve. This is graphically illustrated in figure 6, in the chapter regarding the price determination. The theoretical housing model in equation 2 forms the basis for the modeling of housing prices in MODAG. The empirical modeling assumes that the relations can be approximated by log-linear models. The variables with lower case are on a logarithmic scale.

$$\boldsymbol{pbs - pc = \beta_P + \beta_{P,Y}(rc - pc) + \beta_{P,r} * RRT + \beta_{P,K} k_{83}} \tag{3}$$

Where:

PBS = price index second hand freehold housing
RC = Households real disposable
RRT = real after tax interest rate
K_{83} = housing stock in fixed prices
PC = Private consumption deflator

The real after tax interest rate is defined as:

$$\boldsymbol{RRT = \frac{(1 + RENPF_{300}(1 - TRTMNW))}{\frac{KPI}{KPI_{-1}}} - 1} \tag{4}$$

Where:

$RENPF_{300}$ = Household's average interest on loans from private finance institutions
TRTMNW = average marginal tax on capital income (0,28 after the tax reform of 1992)
KPI = consumer price index

House price in the second hand market

The real housing price long-term sensitivity to changes in real income, real after tax interest rate and housing capital are given by the β_P parameters. The parameters used for (rc − pc), and (pbs − pc) are interpreted as long-term elasticities, while the parameter used for RRT is interpreted as long-term semi elasticity.

The long-term solution for the implemented equation for the housing price in the second hand market:

$$\boldsymbol{pbs - pc = constant - 1,62 * (rc - pc) - 11,59 * RRT - 0,62 * k_{83}} \qquad \textit{(5)}$$

We can see that the real housing price in the second hand market increases with 1% if both the housing capital and the real income increase with 1% each.

Further, we see that the real interest rate has a considerable amount of effect on the housing price. An increase of the real interest rate by 1% reduces the real housing price by over 11%, in a long-term perspective.

5.2 Jacobsen and Naug

In 2004, Jacobsen and Naug (Jacobsen and Naug, 2004:2) published an article which presented an empirical model for predicting housing prices in the Norwegian housing market. The article was presented in The Norwegian Banks journal “Penger & Kreditt” and was named “What drives house prices?”. The authors limited their analysis by describing the housing prices for a given housing stock. The analysis is based on the following aggregated demand function:

$$H^D = f\left(\frac{V}{P}, \frac{V}{HL}, Y, X\right), \qquad f_1 < 0, \quad f_2 < 0, \quad f_3 > 0, \qquad \textbf{(1)}$$

Where

H^D = demand after housing

V = total housing costs for a typical owner

P = price index on goods and services other than housing

HL = total housing costs for a typical tenant (rent)

Y = household’s real disposable income

X = a vector of other fundamental factors affecting housing demand

F_i = the derived function F (*) with respect to argument i

Equation 1, tells us that demand for owner occupied dwellings increases with higher household income, and decreases if the owner costs increases in relation to the costs of a tenant or prices of other goods and services other than housing. The vector X contains observable variables, which capture effects of demographic conditions, banks' lending policies and household expectations concerning future income and housing costs. Expectations concerning future income and housing costs are important because (a) housing is a consumer durable (b) the purchase of a dwelling is the most substantial purchase for most households during their lifetime and (c) most households debt-finance a substantial portion of the purchase when buying their first home or when trading up in the housing market. The housing cost for an owner-occupier measures the value of goods, which the owner-occupier relinquishes by owning and occupying a dwelling for a period. The real housing costs for owners may be defined as:

$$\frac{V}{P} \equiv \frac{PH}{P} BK = \frac{PH}{P}\left[i(1-\tau) - E\pi - \left(E\pi^{PH} - E\pi \right) \right], \qquad \textbf{(2)}$$

Where

BK = housing cost per real krone (NOK) invested in a dwelling

PH = price for an average dwelling (in NOK)

i = nominal interest rate

τ = marginal tax rate on capital income and expenses

Eπ = expected inflation (expected rise in P and HL, 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. The expression [EπPH – Eπ] is the expected real rise in house prices. Expected housing wealth increases if [EπPH – Eπ] increases. This means that the real housing costs for owners fall. Thus, it becomes relatively more advantageous to own a dwelling than to rent, and demand for owner-occupied dwellings rises. The two equations 1 and 2 sums up the demand for owner occupied housing.

The authors tested for effects of the following variables:

- Households' total (nominal) wage income
- Indices for house rent paid and total house rent in the consumer price index (CPI)
- Other parts 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 relocation/centralisation
- TNS Gallup's indicator of households' expectations concerning their own financial situation and the Norwegian economy (the consumer confidence indicator)

This is the following empirical model Jacobsen and Naug concluded with in the article using the least square method with an estimation period from 1990Q2 to 2004Q1. It was not feasible to include all the explanatory factors in one house price equation with a meaningful result.

$$\boldsymbol{\Delta house\ price_t} = \boldsymbol{0,12\Delta income_t - 3,16\Delta(INTEREST(1-\tau))_t - 1,47\Delta(INTEREST(1-\tau))_{t-1} + 0,04EXPEC_t - 0,12[(house\ price_{t-1} + 4,47(INTEREST(1-\tau))_{t-1} + 0,45unemployment - 1,66(income - housingstock)_{t-1}] + 0,56 + 0,04S_1 + 0,02S_2 + 0,03S_3}$$

$R^2 = 0.8773$, $\sigma = 0.014166$, DW = 2.57.

Where:

houseprice = Price index for resale homes

INTEREST = Banks' average lending rate

τ = Marginal tax rate on capital income and expenses (0.28 since 1992)

$EXPEC = (E–F) + 100 \quad (E–F)^3$

E = Indicator of household expectations concerning their own financial situation and the Norwegian economy. Measured as rate, total over two quarters. Source: TNS Gallup

F = Value of E that may be explained by developments in the interest rate and unemployment. Calculated from an estimated model of TNS Gallup's consumer confidence indicator

unemployment =Unemployment rate

income = Total wage income in the economy

housingstock = Housing stock at constant prices

Si = Variable which is equal to 1 in quarter i, otherwise zero.

The variables with small letters are measured on a logarithmic scale.

We now have a basis for further analysis and modelling of the housing market. We will now combine our own believes with previous research, and present factors which we think affect the housing prices.

6. Development in possible factors affecting the Norwegian housing market

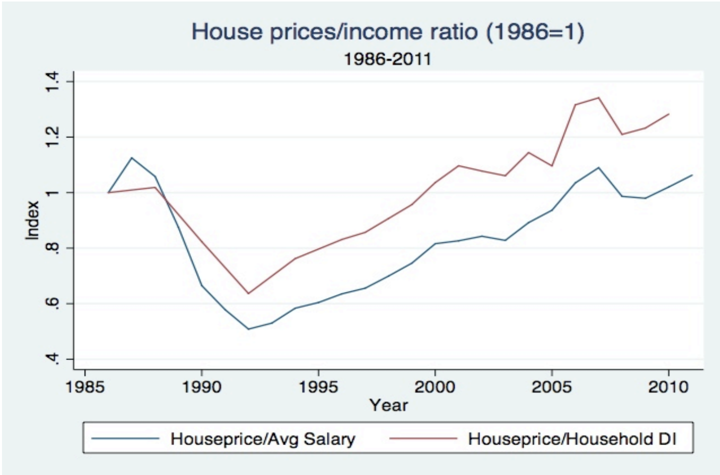
On the basis of previous works and the previous discussed housing models, we have studied and examined the development in Norwegian households income, debt level and expectations, along with the real interest rate, unemployment rate, population growth, building costs and number of traded and initiated dwellings.

6.1 Norwegian households' disposable income

The households' real disposable income has increased considerably since 1970. This factor represents the households' possession of purchasing power. With higher demand for houses, the buyers will push the prices from their willingness to pay towards their ability to pay. This behaviour may contribute to abnormally high prices. We have collected data from the households' income after tax and public benefits, as well as the average annual salary, independent of industry. The average annual salary is calculated by dividing total amount of salary in all industries by the total amount of full-time equivalents in Norway. The salaries reported in the national accounts are defined as salary costs, less employers' social security contributions and pension premiums. This should include regular salaries incurred during the production process, bonus payments related to working conditions, and benefit advantages of free car, free telephone, newspapers etc. It also includes overtime and wages paid to employees during periods of sickness. Social security payments are not considered as salary in the national accounts.

We divided the housing prices with the household's disposable income and the average salary. The ratio is illustrated in the figure 7. This ratio illustrates the development in housing prices compared to income. This ratio is considered as an indication of fundamental value, and a potential deviation will be corrected over time (IMF, 2004). As we can see from the figure 7, today's ratio (average salary) is approximately at the same level as before the de-regulation of the credit market in the mid-1980s, but still under the price-top in 1987. The fall in prices during the banking crisis caused the ratio to halve, but the increase in housing prices after 1993 has been stronger than the increase in income, so the ratio is almost restored at 1987 level.

Figure 7. House prices divided by income. 1986=1



After the financial crisis and the resulting turmoil in international markets, the housing prices have recovered quickly. If we look at the years from 2007 to 2011 as a whole, the development in income has been stronger than the development in housing prices, resulting in a lower ratio in 2011 than in 2007.

6.2 Household's mortgage

Most Norwegian households live in mortgage financed homes. The majority of the Norwegian people live in freehold residential or cooperatives (Eurostat, 2012). The value of the dwelling represents the majority of the household's gross wealth. At the same time, most of the households' debt is mortgage with the dwelling used as collateral.

In reality, debt and housing prices can affect each other in different ways. If we assume that a household finance their home investment with credit, we can also assume that the aggregated credit among households increase correspondingly in the first time sale of a dwelling, since the seller will normally not be another household. An increase in the price of new houses will therefore lead to a higher aggregated credit (Jacobsen and Naug, 2004:2).

Another way the debt and housing prices affect each other is by first-time purchases and last-time sales. When a new household enters the market, another household will necessarily exit the market. This household frees up capital, which can be used for mortgage down payments. If the down payment is less than the first-time buyer's loan, the aggregated credit will increase. The opposite if the down payment is larger. An increase in prices of used houses will therefore lead to a higher aggregated credit among households. This is because the first-time buyer needs more credit to finance the purchase, and the seller's remaining mortgage will probably be smaller than the first-time buyers loan (Jacobsen and Naug, 2004:2). Another situation is when used houses are traded between households where no one enters or exits the market. If we assume the prices are constant and one household buys a more expensive house, and the other household buys a correspondingly cheaper house, the aggregated credit will remain unchanged if the freed capital is used for mortgage down payment, and if the household which is buying the more expensive house, is credit financing the difference (Jacobsen and Naug, 2004:2). An increase in prices in this example will result in the acquiring household needing more credit to finance the price difference. The aggregated credit among households will therefore increase (Jacobsen and Naug, 2004:2). Increased housing prices can lead to a long period with high debt growth. Some houses will be traded during this period, and the aggregated credit increases. Gradually, the prices will stabilize, but for a long time houses will be traded for higher price than the last time the house was traded. In theory, the increase in prices will contribute to credit growth until the entire housing stock is traded at the new price level.

If we take a closer look at the ratio between the households' gross debt and the total value of the housing stock, illustrated in figure 8, the ratio was decreasing towards the abandoning of the regulation. After many years with regulation, the government stopped controlling the

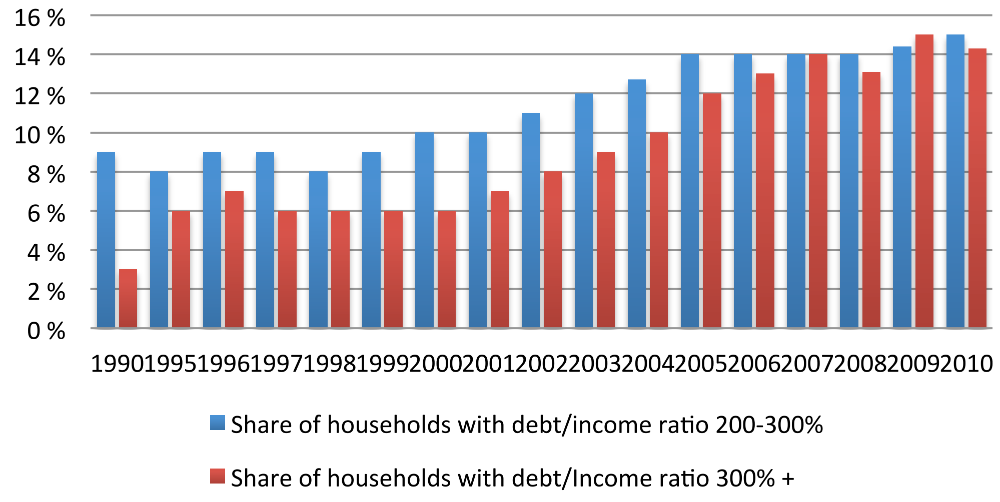
Figure 8. Households' total gross debt divided by total housing stock value. 1970-2010. 2007=1. (Jansen, 2011)

banks lending to households. The total outstanding debt increased more than the total value of the housing stock. During the banking crisis and the drop in housing prices towards 1993, the ratio increased to the double of the ratio level from 1982. Since 1993, the total value of the housing stock has increased much more than the household's debt. In the years after

2000, the ratio has remained relatively stable, implying that the debt has increased along with the housing prices. Although, today's ratio is significantly lower than during the banking crisis. The increase in housing prices has also had a reinforcing effect on the aggregated credit among households. The households can put a greater property value as collateral for their mortgage, which gives incentives for increased borrowing. Especially since the banks have small losses related to mortgages (Gjedrem, 2010). The credit growth flattened during 2011 (SSB, 2012:1) and it was especially the development in other types of debt that mostly contributed to the decline. According to the Norwegian central bank's lending survey from the 4th quarter 2011, it was the banks' tightening of lending policy in the beginning of 2011 that led to the decline. On the other hand, the credit growth in mortgages with the property as collateral is still increasing (SSB, 2012:1).

Another interesting ratio to examine is the debt as a share of the household's disposable income. This ratio has increased over time, especially in the last decade. The total gross debt represented more than 200% of households' total disposable income in 2010 (Jansen, 2011). In other words, the households have twice the amount in debt as in disposable income on average. This is furthermore illustrated in figure 9, where we can see that the share of households with debt/income ratios over 200% has more than doubled since 1990.

Figure 9. Percentage of households by debt/income ratio



The increased housing prices may have given the households more collateral when refinancing their homes and many households have experienced very good return on their home investment.

Anundsen and Jansen’s paper from 2011 (Anundsen and Jansen, 2011) clearly indicate that the relationship between housing prices and credit is mutually self-reinforcing. First, a higher income leads to increased property valuations, which raises the value of the collateral. This spills over to an increased credit supply, stimulating housing prices further and so on. Anundsen and Jansen (2011) show that an increase in housing prices by 1% results in an increase in household debt by 0,76%. It also concludes that the housing prices and the households’ debt have a relationship with an elasticity of 0,98. Correspondingly, it shows that an exogenous shock in the credit aggregate by 1% results in an immediate change in housing price growth rate by 0,86%.

6.3 The real interest rate

A change in the interest rate changes the capital cost of owning and hence the cost of living. Lower cost of owning allows households to afford more expensive dwellings. A low interest rate can therefore contribute to an increase in the housing prices. The low interest level before and after the financial crisis has probably stimulated the housing demand and consequently the prices (Grytten, 2009) (SSB, 2012:1).

The current low interest level has resulted in low interest costs among households. However, the interest cost among many Norwegian households will increase sharply if the interest level rises. The high debt level among Norwegian households has increased their vulnerability to a decrease in income, unemployment and increasing interest rates. Since the debt is unevenly distributed among households, estimates show that a large share of the households will get an interest burden of more than 30% of disposable income if the interest rates are normalized (The Financial Supervisory Authority of Norway, 2011). The growth in debt among households has been higher than the growth in income for many years. Aggregated debt was in 2011 twice the amount of total disposable income. Since the largest share of mortgages has a floating interest rate, an increase in interest will immediately affect household consumption. Calculations performed by “Statistics Norway” and “The financial supervisory authority of Norway” shows that 20% of Norwegian households will have an interest burden considered as high if the interest rates are normalized (The Financial Supervisory Authority of Norway, 2011).

The banks’ lending policy depends on the bank’s profitability on lending, government regulations and the customers’ expected solvency and collateral. The real interest rate is also dependent on interest deductibility determined by the government. Before 1992 there was full interest deductibility, but the tax reform implemented in 1992 reduced the deductibility from the taxpayer's marginal tax rate to 28% flat. One would probably think such a change would have a negative impact on the housing prices, but the housing market turned around at the time the reform was implemented. Studies conducted on the outcome of the reform could not conclude that it had a significant effect on housing prices (Sommervoll, 2007). The explanatory factors for the strong increase in prices were the already low price level, low inflation and low interest rates, rather than the tax reform.

The real interest rate is calculated by adjusting the banks' average interest rates on loans to households, tax savings and inflation. The average interest on banks' lending was gathered from SSB and the Norwegian Bank. The statistics are weighted average interests, included provisions on credit loans secured with the loaner's house. The banks' average interest rate is calculated with the following formula:

$$\frac{(Interest\ Loan\ 1 * Amount\ Loan\ 1 + Interest\ Loan\ 2 * Amount\ Loan\ 2)}{(Amount\ Loan\ 1 + Amount\ Loan\ 2)}$$

Then the average among banks is calculated with the following formula:

$$\frac{(Interest\ Bank\ 1 * Amount\ Bank\ 1 + Interest\ Bank\ 2 * Amount\ Bank\ 2)}{(Amount\ Bank\ 1 + Amount\ Bank\ 2)}$$

The Norwegian tax system favours a housing investment in several ways.

- There is a very low property tax, and the property tax varies between 0,2 and 0,7% of property's market value
- Income from renting out a part of your own house is exempted from tax[3]
- Interest costs associated with debt are tax deductible by 28%
- Taxable value of primary housing is maximum 30% of market value, while liabilities are 100% of market value

[3] The owner is required to live in the house, and the rental price must not exceed 50% of the property's total rental value.

The inflation in the economy can be calculated in several ways. The most familiar ones are the “consumer price index” (CPI), the “gross domestic product –deflator” (GDP deflator) and the “private consumption deflator in the national accounts (PC). The CPI measures the inflation in prices on products and services purchased by households. The percentage change in the CPI is often used as a measure on the total inflation in the economy. The CPI is used in wage negotiations, adjustments of private rental contracts, business agreements and as a deflator in the national accounts when calculating fixed prices for comparisons. The estimation of the CPI is performed monthly by SSB in Norway. The different prices are weighted by the results of the consumer survey, which maps the household’s consumption. Every month prices are gathered from 1 700 businesses and the data material includes 40-45 000 observations from a selection of 800 products and services.

In economics, the GDP deflator is a measure of the level of prices of all new, domestically produced, final goods and services in an economy. Unlike the CPI, which reflects the prices of a representative basket of goods and services purchased by the consumers, the GDP deflator measures the inflation in the entire domestic economy and compares the price of currently produced goods relative to price of goods in the base year. Therefore, changes in the price of imported goods affect the CPI, but not the GDP deflator. Also, changes in the price of domestically produced capital goods affect the GDP deflator, but not the CPI. Changes in the price of domestically produced consumer goods are likely to affect the CPI more than the GDP deflator because it is likely that those goods make up a larger part of consumer budgets than of the GDP. Price changes may cause consumers to switch from buying one good to another. Whereas the fixed-basket CPI does not account for altered spending habits caused by price changes, the PC deflator's ability to account for such substitutions makes it a much used measure of inflation as well.

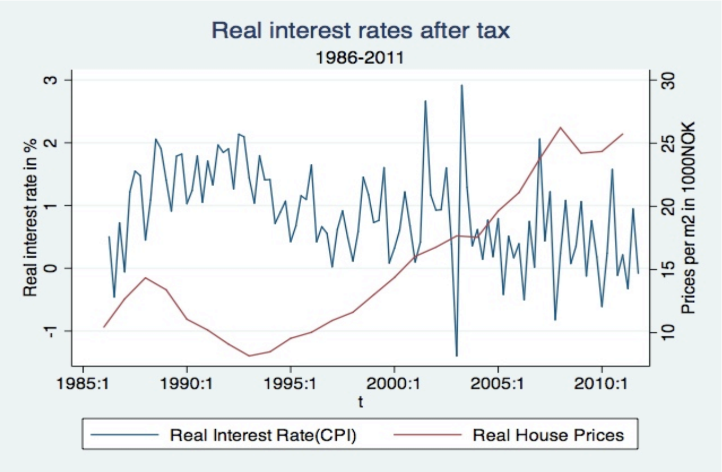
The GDP deflator has increased by 4,8% on average every year since 2001. At the same time, the CPI has only increased by 1,8% on average per year (SSB, 2012:3). This is not common if we look at a historical perspective. The scenario was the opposite just a decade ago, when the CPI increased more than the prices of domestically produced goods and services. The main reason for this has been the increased access to cheap Asian goods as well as the export prices, especially oil and gas, has exploded (Bjørnland, 2012). In addition, this will help strengthen the Krone, which makes it even cheaper to import consumer goods. It will drive consumer prices further down and at the same time help keeping the Norwegian central bank’s key policy rate at bay

(Bjørnland, 2012). The interest rate deflated by CPI has been more stable and at a higher level, with some exceptions, than the GDP deflated interest rate. This can be interpreted as Norway getting paid more and more for the goods they produce and export, while they pay less and less for the goods they import. We have calculated the real interest rate after tax based on CPI using the following equation (Ministry of Justice and Public Security, 1994):

$$RIRAT = \frac{Nominal\ interest * (1 - marginal\ tax\ rate) - inflation}{(1 + inflation)}$$

This is illustrated in figure 10:

Figure 10. Real interest rates after tax deflated by CPI and real housing prices (SSB)



6.4 Consumer expectations

One factor that could help explain the development in the housing market is the consumer expectations. This is because psychology is important to make assumptions of the households' economic behaviour. If people expect good outlooks for the economy, both in general and for their personal economy, they can become more willing to borrow and spend money. Many expectations can also be self-fulfilling. An easy example is expectations for rising house prices, making the demand for houses higher (people want to buy now instead of later at higher prices). Therefore, the expectations may contribute to higher prices.

The Norwegian Trend Indicator (Norsk Trendindikator) made by TNS Gallup is a good measurement for consumer expectations in Norway. This indicator measures expected future demand from consumers. Evaluation of consumer confidence has been examined together with the following demand the last 60 years by the University of Michigan in the US, and is found to be relevant. The Norwegian Trend Indicator has three main elements (TNS Gallup, 2012):

- The respondents perception of personal economy
- The respondents perception of the nation's economic outlook
- The respondents evaluation of whether or not now (today) is a good time to make larger personal investments

The survey has been performed four times a year since the third quarter in 1992 (FNO, 2012). The calculation of the indicator summarizes the differences between optimistic and pessimistic answers, and calculates the average response. These are the published indicator values. The main indicator equals the difference between the percentage of positive and negative answers for each question, divided by five. The main indicator is also adjusted for trend and seasonal effects, to clarify the development in consumer expectations. This means that the indicator will be negative when the consumers have negative expectations for the economy in the close future, around zero if they believe it is going to be status quo, and positive when the respondents have positive expectations for the economy.

Figure 11. Norwegian trend indicator and housing prices, (SSB:3) (TNS Gallup).

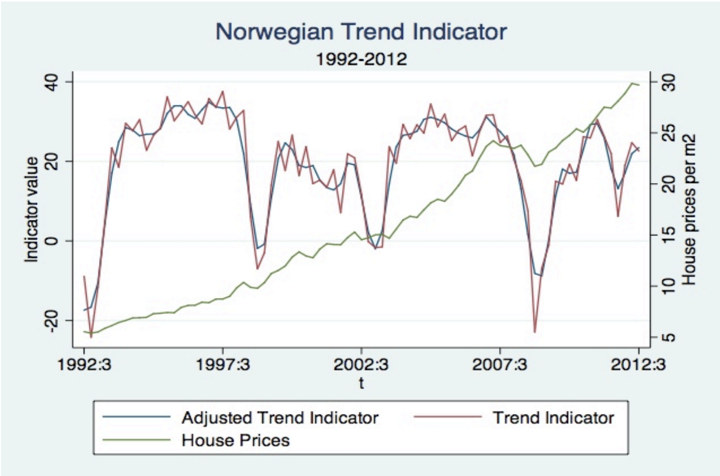


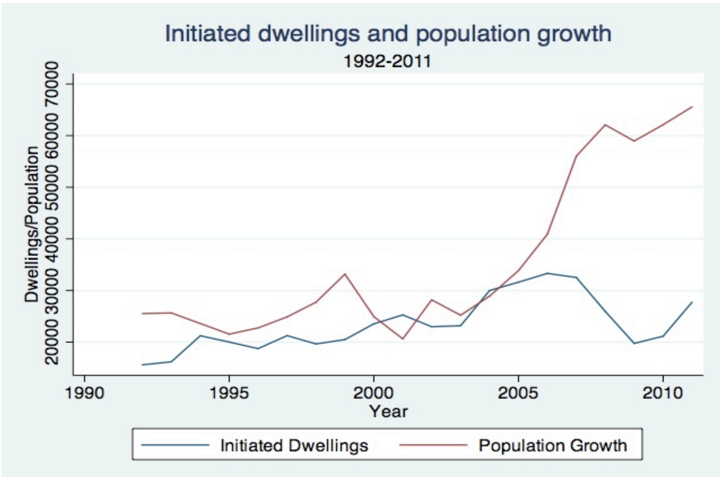
Figure 11, illustrates the adjusted Norwegian Trend Indicator and non-adjusted trend indicator, compared with the average house prices per square meter in Norway. The Trend Indicator is decreasing in the beginning of the chart, caused by the recession in the years before 1993. We can see from the chart that the confidence is at a positive level except in the periods of the three major crises mentioned earlier in the thesis. It is more difficult to see the impacts on the housing prices, but you can see minor falls or stagnations in all the crises mentioned above. It does not mean that the consumer confidence is the main reason for the fluctuations in the housing prices, but we can see some coincidental movements.

6.5 The number of initiated dwellings

One way to examine the development in the housing market is to look at the number of initiated houses and apartments. When the market cannot supply enough dwellings to clear the market, the prices will rise in the short run. This will again increase the number of initiated dwellings, since contractors will find it more profitable to start new projects. A high number of initiated dwellings will therefore imply an excess demand in the market. In the long run, initiated dwellings will increase the supply and might have a dampening affect on prices.

Similarly, the housing turnover might indicate something about the state in the housing market. A higher turnover will probably increase the pressure on prices. Many people changing houses or wanting to get into the market will increase the demand, and hence the prices will increase. Theoretically this will lead to increased building of new houses as well. Of course there are situations where a higher turnover could be the result of a lot of bankruptcies and payment problems, which will lead to forced sales of houses, and hence a higher turnover can imply decreasing prices. This would be a very rare situation, at least in Norway. Most people may like to keep their houses and keep living in the same dwelling when prices fall instead of selling with a loss, and rather sell when the market turns.

Figure 12. Number of initiated dwellings and population growth (SSB:3)



We have shown the number of initiated dwellings and the population growth in the figure 12. The population growth in numbers is calculated as the number of new-borns minus the number of dead, plus net immigration (immigrants minus emigrants).

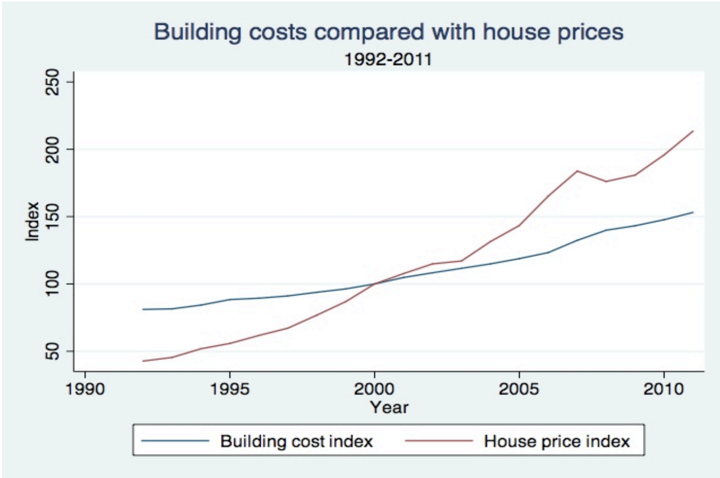
Because of the increasing population growth, the building of new houses has mainly increased since 1992. But as we can see, the level of initiated dwellings is lower than before the financial crisis in 2008. Since we know that the housing prices have continued to grow after the crisis, it shows that the initiated dwellings are affected by other factors than just the housing prices. This is natural when we know that this industry is highly cyclical. We can see that the numbers of new dwellings are increasing at approximately the same rate as the population growth until 2005. After 2005, and during the financial crisis in 2008 there seem to be a mismatch between the two, which could be a part of the explanation for the increasing housing prices. Norway did fairly well through the crisis, and the immigration remained high because of the high level of employment. The contractors were affected by the turmoil in the international markets, and experienced difficulties making large investments in new projects. This might have created the gap shown in figure 12, and may have contributed to increasing demand in the Norwegian housing market.

6.6 Building costs

There are two main factors affecting the costs of building new dwellings: building costs and the price of land. The price of land is hard to measure, and is mostly reflected in the price development for housing in general. For the building costs we refer to the building cost index, which has been developed since 1972 by Statistics Norway.

The index measures the price development in building materials, machinery, transportation and other factors included in house construction. The index accounts for changes in the price of input factors, as well as their relative shares of total costs. All goods and services used in the production of houses are included in the index, and the prices are measured each month. The relative weights are measured every 10 years (SSB, 2010). An important element to remember is that this index does not include any changes in productivity or the profit margins of the entrepreneurs and contractors.

Figure 13. House price index and building cost index. 2000=100 (SSB:3)



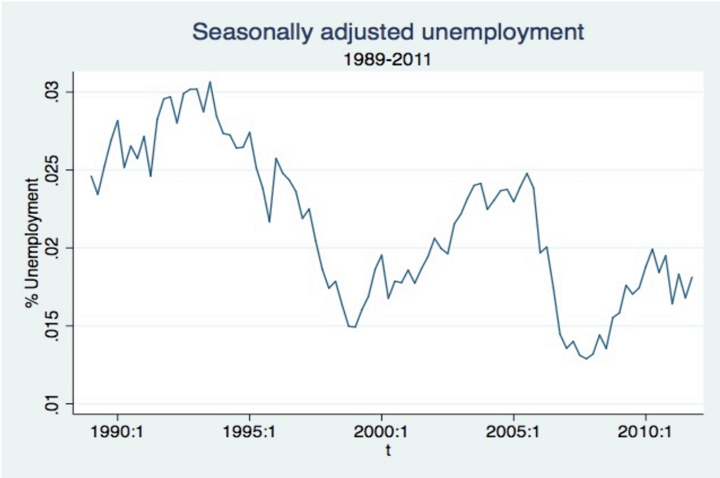
As we can see from figure 13, the house prices have increased significantly compared to the building costs since 1992. Since the figure is illustrating index values, the only information we observe is that the house prices has increased a lot more than the building costs since 1992.

6.7 Unemployment

The development in the labour market is important for households' perception of future income and solvency. In Norway, a relatively large share of the population is employed, compared to other countries. One of the reasons is the large share of employed women. Low unemployment levels can contribute to a better and more secure economic outlook and income growth. We can assume that a low unemployment rate can lower the threshold for credit financing of consumer goods, dwellings as well. The level of employment is closely related to the income level in the economy, since the wages in the economy are mainly negotiated by the employers' organizations and workers' unions. The respective bargaining power of the unions depends on the current situation in the labour market. In the situation where the unemployment rate is low, the real wages will increase due to the high demand for labour. High demand will result in workers having more employment options.

An increase in real wages will result in the employers demanding less labour. Low wages and high unemployment can contribute to a decline in borrowings and consumption among households.

Figure 14. Seasonally adjusted unemployment rate (SSB)



As we can see from figure 14, the unemployment rate in Norway increased as a result of the financial crisis. Since 2010 the unemployment rate has decreased. Some industries are still struggling after the financial crisis and are experiencing lower employment than the pre-crisis level.

We now have a better understanding of the development of the different factors and their influence on the housing prices. Before we start analysing these factors statistically, we will present the statistical theory and methods we have used in our thesis.

7. Statistical theory and methods

In this chapter we will present the different methodology and models we will be using in our thesis. First, we will explain the basic regression and the least square method. We will also present several important elements and tests related to the work with time series data, and the consequences each of them may imply. Then we will explain the HP-filter and how we will later use it for estimating trends.

7.1 Multiple regression analysis

Regression analysis is a statistical method to evaluate the relationship between independent variables and a dependent variable. We use multiple regression analysis to test whether it exists statistical relationships between the independent variables and the dependent variable. It is important to remember that we can reveal statistical relationships that do not comply with economic reasoning and theory. When two variables have no direct causal connection, but we wrongly infer that they do, it is called spurious regression. More specifically, regression analysis helps one understand how the typical value of the dependent variable changes when any one of the independent variables is varied, while the other independent variables are held fixed. The choice of independent variables should be theoretically grounded. For example, a variable can prove to be significant in affecting the housing prices, but it does not prove that this affect the housing prices in reality. We can never be certain about the underlying causal mechanism. There exist several statistical assumptions regarding the multiple regression analysis:

1. The error term is a random variable with a mean of zero, $E(u_t) = 0$
2. The error term is normally distributed, $u_t \sim N(\mu, \sigma^2)$
3. The errors are linearly independent of each other $cov(u_i, u_j) = 0$
4. The variance of the errors is constant and finite over all values of x_i, $var(u_t) = \sigma^2$

   It is customary to divide this assumption into two parts:

   - Homoscedasticity – the error term has the same variance in each observation
   - Non-autocorrelation – the errors are uncorrelated between observations

5. The independent variables are not correlated with the error term, $cov(x_t, u_t) = 0$
6. A linear relationship between the dependent variable, independent variables and the error term $y_t = \beta_0 + \sum_{i=1}^{i} \beta_i x_{i,t} + u_t$

7.2 Least square method

To find the equation that best fits the observations, we use the least squares method. This method is mathematical and estimates the relationships that minimize the total squared deviations between the actual and predicted values. The method easily generalizes to finding the best fit of the form

$$y = a_1 \mathrm{f}_1(x) + \cdots + a_k f_k(x)$$

It is not necessary for the functions f_k to be linear in x – all that is needed is that y is to be a linear in parameters. The equation below is with just one variable for easy intuition. We minimize the sum of squared errors (SSE):

$$SSE = \sum_{i=1}^{n} \left(Y_i - \hat{Y}\right)^2 = \sum_{i=1}^{n} \left(Y_i - (b_0 + b_1 X_i)\right)^2$$

The solution is: $\hat{b}_1 = \frac{\sum_{i=1}^{n} X_i - \bar{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = \frac{cov(X,Y)}{var(X)}$ and $\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}$

7.3 Analysis of time series

When analysing time series, there are several important considerations to be made in order to get unbiased and consistent estimates. Many of the tests presented are complicated and rarely calculated manually, but we will try to explain them intuitively. In this section we will explain which matters to consider.

7.3.1 Adjusting for seasonality

Seasonality refers to particular time frames when time series are subjected to, and influenced by, recurring tendencies that produce patterns that are apparent in the analysis of the data. By adjusting for seasonality, we can analyse the fluctuations in the series, independently of the recurring tendencies. The fluctuations in the adjusted series will therefore be explained by other factors than seasonality.

We will assume that the series consist by the components long-term trend (L), cycle (C), seasonal effects (S) and noise (I) with a multiplicative relationship (Klovland, Helliesen and Kvåle, 2012). That gives us the following equation: $Y_t = L_t \times C_t \times S_t \times I_t$. We then want to remove the seasonal component, and isolate $Y_t^{LCI} = L \times C \times I$.

To adjust for seasons in quarterly data, we go through five steps:

1. Isolate variation from trend and cycle $L \times C$
    a. We define $Y_t^{LC} = \frac{1}{4}(Y_{t+2} + Y_{t+1} + Y_t + Y_{t-1})$ so we have a rolling average for years of data.
    b. We then assume that variation from seasonal effects and noise are mainly smoothed out, so that $Y_t^{LC} = L \times C$.
2. Isolate variation from seasonal effects and noise $S \times I$
    a. We can now define $Y_t^{SI} = \frac{Y_t}{Y_t^{LC}} = \frac{L \times C \times S \times I}{L \times C} = S \times I$
    b. Y_t^{SI} will then be an estimate on variation from seasonal effects and noise.
3. Isolate the variation from seasonal effects S
    a. We want to isolate the seasonal component in Y_t^{SI}.
    b. We then define $S_q = \frac{1}{N_q}\left(Y_{1,q}^{SI} + Y_{2,q}^{SI} + \cdots + Y_{N_q,q}^{SI}\right)$; where N_q is the number of years with quarterly (q) observations.
        i. We are smoothing out noise by taking the average of every quarterly observation of Y_t^{SI}
        ii. S_q will then be an estimate on the quarterly effect, constant over years.
4. Normalize the variation from seasonal effects S^*
    a. We want to find a normalized index $S_q^* = S_q / \frac{S_1+S_2+S_3+S_4}{4}$

   b. The way S_q^* is constructed $\sum_1^4 S_q^*$ will always be 4.
5. Isolate the variation from trend, cycle and noise $L \times C \times I$
   a. We have normalized a seasonal index S_q^*.
   b. Define $Y_t^{LCI} = \frac{Y_t}{S_q^*}$; which means that all observations will be divided with corresponding normalized seasonal index
   c. Y_t^{LCI} will then be our estimate on variation from trend, cycle and noise.

This method could easily be adjusted to account for monthly data. It is also possible to use dummy variables, to try extracting the seasonal effects.

7.3.2 Stationary data

A time series is weakly stationary if the series' mean, autocorrelation and variance are constant over time. Most business and economic time series are far from stationary when expressed in their original units of measurement, and even after deflation or seasonal adjustment they will typically still exhibit trends, cycles, random-walking, and other non-stationary behaviour (Duke Education, 2005). Economic time series often have a long-run trend, seasonality and inflation affecting the data. For a stationary series, 'shocks' will gradually fade away. The opposite will happen with non-stationary data, where the persistence of shocks will sustain. So that for a non-stationary series, the effect of a shock during time *t* will not have a smaller effect in time *t+1*, and in time *t+2* etc. (Brooks, 2008). If the time series tend to revert to the trend line after a disturbance, it is possible to stationarize it by de-trending, logging and deflation of the time series (if I(0)).

The trend is in most cases stochastic or deterministic. A time series with a deterministic trend has a mean that grows around a fixed trend, which is constant and independent of time. A time series with a stochastic trend is a non mean-reverting process that can move away from the mean either in a positive or negative direction. It follows a pattern, which is affected by past-observed values (t_{-1}).

We can transform the time series with a deterministic trend to be stationary, by estimating the trend and subtracting it before we use the data for modelling. If the time series has a stochastic trend, we can make it stationary by differencing the data (subtracting y_{t-1} from y_t). However, this will not work if the variable is integrated by a higher order than 1(this is explained on the next page). Sometimes the non-stationary series may consist of a stochastic and deterministic trend at the same time.

Testing the time series for non-stationarity

To test whether a variable is stationary or not, we can use graphical analysis. Studying graphical illustrations of the data can give us an indication of what we can expect from the data, but it cannot replace statistical tests. We have used a test called the augmented Dickey-Fuller test.

Dickey-Fuller test

We can test if a variable is stationary by using the Dickey-Fuller test. This is a unit root test. If we have a non-stationary variable, we want to find out which order it is integrated by. If a non-stationary series, Y_t must be differenced *d* times before it becomes stationary, then it is said to be integrated of order d. This would be written Y_t I(d) (Brooks, 2008). A non-stationary variable that becomes stationary after one differencing is denoted by I(1) - integrated by first order. The first step is to difference the variable by one period. We have three different types of non-stationary processes:

1. Test for unit root ("Random walk")
$$X_t - X_{t-1} = (\rho - 1)X_{t-1} + \varepsilon_t = \delta X_{t-1} + \varepsilon_t$$
2. Test for unit root with drift
$$X_t - X_{t-1} = \mu + (\rho - 1)X_{t-1} + \varepsilon_t = \mu + \delta X_{t-1} + \varepsilon_t$$
3. Test for unit root with drift and a deterministic trend
$$X_t - X_{t-1} = \mu + (\rho - 1)X_{t-1} + \lambda T + \varepsilon_t = \mu + \delta X_{t-1} + \lambda T + \varepsilon_t$$

Let us say that the variable in question is a non-stationary variable characterized by "random walk with drift and a deterministic trend

$$X_t = \mu + \rho X_{t-1} + \lambda T + \varepsilon_t$$

$$X_t - X_{t-1} = \mu + (\rho - 1)X_{t-1} + \lambda T + \varepsilon_t = \mu + \delta X_{t-1} + \lambda T + \varepsilon_t$$

Where μ is the drift, T is the deterministic trend and ε is the error term.

The null hypothesis in this model is will be H_0: $\delta = 0$, implying non-stationarity. Our alternative hypothesis is that $\delta < 0$. If the null hypothesis is rejected the variable is stationary, I(0). If we cannot reject the null hypothesis, we have to perform a new test where H_0 says that the variable is integrated by second order, I(2), and the alternative hypothesis says that the variable is integrated by first order I(1). Thus, testing H_0: ΔX_t I(1) is equivalent to H_0: of X_t I(2). The rejection regions for the null-hypothesis are to find in special Dickey-Fuller tables. All the three tests have different rejection regions, depending on the model used and the number of observations. A weakness in the model is that it has difficulty distinguishing between $\delta = 0$ and when δ is close to zero. Note that testing with zero intercept/drift is extremely restrictive, so much so that it is hard to imagine ever using it with economic time series (Davidson and MacKinnon (1993), referred in Elder and Kennedy (2001)).

The Dickey-Fuller test assumes that the error term is white noise. White noise implies that the error term is uncorrelated with the error term in other periods in the time series. If this is not the case, we can perform an extended Dickey-Fuller test, called the augmented Dickey-Fuller test (ADF). The ADF-test adds lagged variables of ΔX_t, in order to transform the error term into white noise. The model captures the correlation between the dependent variable and previous time periods in a better way. The test can be expressed as:

$$\Delta X_t = \mu + \delta X_{t-1} + \lambda T + \gamma_1 \Delta X_{t-1} + \gamma_2 \Delta X_{t-2} \ldots + \gamma_n \Delta X_{n-1} + \varepsilon_t$$

Now a new problem arises, how many lags are optimal? Two simple rules of thumb are suggested by Brooks (2008). First, the frequency of the data can be used to decide. If the data are quarterly, as in our case, we can use 4 lags. Second, an information criterion can be used to decide. We then choose the number of lags that minimises the value of an information criterion. The information criteria consists of two factors, the residual sum of squares and the penalty resulting in the loss of degrees of freedom when adding an independent variable. If we include many lags, the model would probably fit better, but a model with few lags are a lot simpler to use and understand. So, when adding a new independent variable we will experience two competing effects on the information criteria: the residual sum of squares will fall, but the value of the penalty term will increase. The univariate criteria could be applied separately to each equation, but it is usually preferable to require the number of lags to be the same for each equation. This requires the use of multivariate versions of the information criteria, which can be defined as (Brooks, 2008):

$$MAIC = \log[\hat{\Sigma}] + \frac{2k'}{T}$$

$$MSBIC = \log[\hat{\Sigma}] + \frac{k'}{T}\log(T)$$

$$MHQIC = \log[\hat{\Sigma}] + \frac{2k'}{T}\log(log\ (T))$$

Where $\hat{\Sigma}$ is the variance-covariance matrix of residuals, T is the number of observations and $k'$is the total number of independent variables in all equations, which will be equal to $p^2k + p$ for p equations in the VAR system[4], each with k lags of the p variables, plus a constant term in each equation. The values of the information criteria are estimated for 0, 1, ..., k‘ lags. The chosen number of lags is that number minimizing the value of the given information criterion (Brooks, 2008).

In most cases, the conclusion will not be qualitatively altered by small changes in the number of lags, but sometimes it will. Including too few lags will not remove all of the autocorrelation, but it will bias the results. While using too many will increase the standard errors in the coefficient. The last mentioned effect arises since an increase in the number of parameters reduces degrees of freedom. Therefore, the absolute values of the test statistics will be reduced. This will result in a reduction in the power of the test, implying that for a stationary process the null hypothesis of a unit root will be rejected less frequently than would otherwise have been the case (Brooks, 2008).

7.3.3 Autocorrelation

Autocorrelation, also called series correlation, is one of the major problems in time series econometrics. Autocorrelation means that the error term is correlated between different time

[4]“A Vector Autoregressive model (VAR) is simply an autoregressive process for a vector of variables.” (Beckert, 2011).

periods in the time series. This results in the estimated coefficients no longer being efficient when using least squares method, and we will necessarily not find the model that minimizes the deviation.

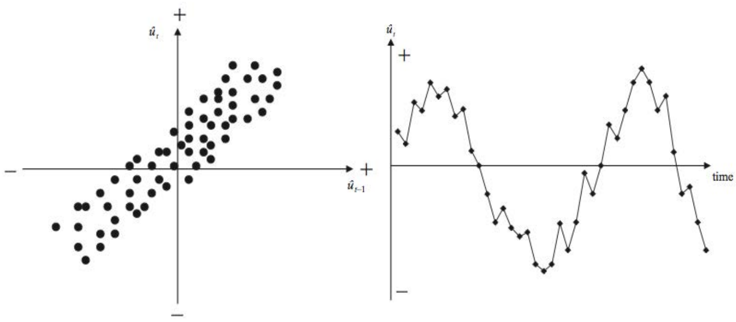
When the deviation is not correct, we can end up taking the wrong decision regarding the relevance of a variable. For example, the least squares method predicts too small deviation under the presence of positive autocorrelation, which can result in an incorrect rejection of the null-hypothesis.

We can test the presence of autocorrelation in the model by examine graphical residual plots or perform statistical test.

Graphical tests

We can say our model contains autocorrelation if the residual plot shows a clear pattern. In order to test for autocorrelation, it is necessary to investigate whether any relationships exist between the current value of $\hat{u}$, $\hat{u}_t$, and any of its previous values $\hat{u}_{t-1}, \hat{u}_{t-2}$ etc. Positive autocorrelation will result in a cyclical residual plot over time.

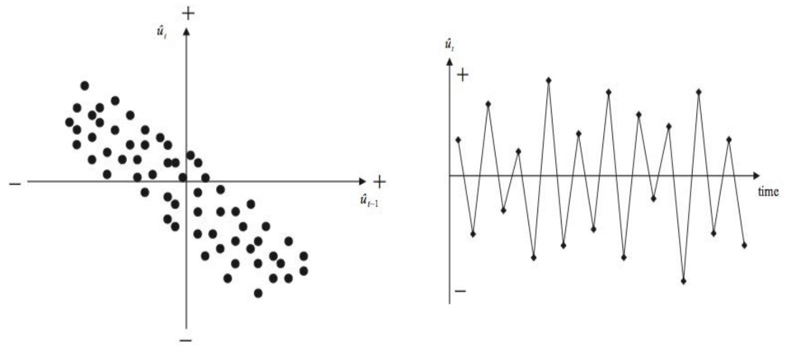
Figure 15. Positive auto-correlation



On average if the residual at time t_{-1} is positive, the residual at time t is likely to be also positive. Similarly, if the residual at t_{-1} is negative, the residual at t is also likely to be negative. Figure 15, shows that a positively autocorrelated series of residuals will not cross the time axis very frequently.

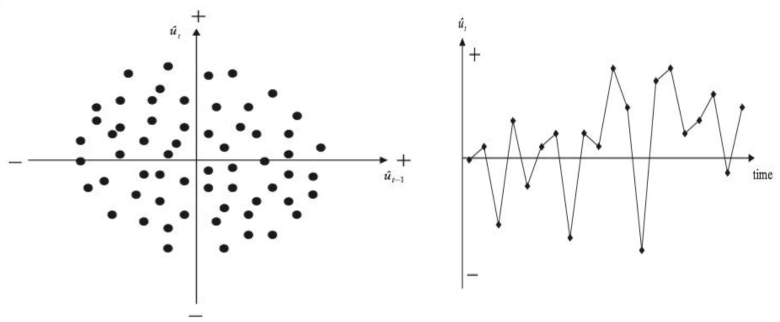
Negative autocorrelation will result in an fluctuating pattern in the residuals. On average if the residual at time t_{-1} is positive, the residual at time t is likely to be negative. Similarly, if the residual at t_{-1} is negative, the residual at t is likely to be positive. Figure 16, shows that a negatively autocorrelated series of residuals will cross the time-axis more frequently than if they were distributed randomly.

Figure 16. Negative auto-correlation



Finally, we have the example where we find no pattern, which points out evidence of autocorrelation. The points are randomly spread and the residuals in figure 17 does not cross the x-axis either too frequently or too little
(Brooks, 2008).

Figure 17. No auto-correlation



Statistical tests for autocorrelation

To examine and interpret graphical plots of variables can be difficult. We should therefore perform statistical tests in addition to the graphical tests. The two tests we will present are the Ljung-Box test and the Breusch-Godfrey lagrange multiplier test for serial correlation.

Ljung-Box test

A test we can use for testing possible autocorrelation is the Ljung-Box Q statistic. This is a combined test, which also tests for higher orders of autocorrelation. The null hypothesis states that there is no evidence of autocorrelation. The Q statistic is chi-squared distributed with *m* degrees of freedom.

$$Q = T(T+2)\sum_{k=1}^{m}\frac{\hat{r}_k^2}{T-k}$$

T indicates the sample size, *m* is the maximal lag length and $\hat{r}_k$ is the value of the estimated autocorrelation coefficient for a given number of lags, *k*.

The Ljung-Box test is a variant of the Box-Pierce with better small sample properties. For $T \to \infty$, the two methods are equivalent.

Breusch–Godfrey serial correlation Lagrange multiplier test

This test allows examination of the relationship between $\hat{u}_t$, and several of its lagged values at the same time. The test is a more general test for autocorrelation up to the *r*th order. The model for the errors under this test:

$$u_t = p_1 u_{t-1} + p_2 u_{t-2} + \ldots + p_r u_{t-r} + v_1, \qquad v \sim N(0, \sigma_v^2)$$

The null hypothesis is that the current error is not related to any of its r previous values. $H_0\colon p_1 = 0, p_2 = 0 \ldots p_r = 0$

$$H_1\colon p_1 \neq 0, p_2 = 0 \ldots p_r \neq 0$$

First, the test obtains the residuals from the least squares method used to estimate the underlying regression. Then the test regresses $\hat{u}_t$ on all the independent variables from the previous regression plus the t-r lagged residuals.

$$\hat{u}_t = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + p_1 \hat{u}_{t-1} + p_2 \hat{u}_{t-2} + \ldots + p_r \hat{u}_{t-r} + v_1, \qquad v \sim N(0, \sigma_v^2)$$

This regression's R^2 is used in the following test statistic:

$$(\mathrm{T} - \mathrm{r})R^2 \sim \chi_r$$

If the test statistic exceeds the critical value from the Chi-squared statistical tables, we reject the null hypothesis of no autocorrelation. One potential difficulty with the Breusch-Godfrey test is the determination of the lag length, r. There is no obvious answer to this, so it is typical to experiment with a range of values, and also to use the frequency of the data to decide. If the data is quarterly, set r equal to 4. If the model is statistically valid, no evidence of autocorrelation should be found in the residuals whatever value of r is chosen (Brooks, 2008).

7.3.4 Normality

We need to test if the normality assumption ($u_t \sim N(0, \sigma^2)$) is violated in order to conduct single or joint hypothesis tests of the model parameters.

One of the most commonly used tests for normality is the Jarque-Bera test. The test checks for both *skewness* and *kurtosis* in the residuals. *Skewness* measures how asymmetrical the residuals are about its mean value and *kurtosis* measures how fat the tails of the distribution are. A normal distribution is characterized by no skewness and a kurtosis coefficient of 3(*mesokurtic*) (Brooks, 2008). The Jarque-bera test checks whether the coefficients for skewness and excess kurtosis are jointly zero, which can be described respectively as:

$$b_1 = \frac{E[u^3]}{(\sigma^2)^{\frac{3}{2}}} \qquad\qquad b_2 = \frac{E[u^4]}{(\sigma^2)^2}$$

$E[u]$ is the estimate for the central movement and σ^2 is the variance. Since the kurtosis of a normal distribution is three, the excess kurtosis is given by (b_2-3). The test statistic is written as:

$$W = T[\frac{b_1}{6} + \frac{(b_2-3)^2}{24}]$$

Where T is the sample size. The test asymptotically follows a $\chi^2(2)$. The null hypothesis is that the data series is both symmetric and mesokurtic.

7.3.5 Heteroscedasticity

The assumption of the classic linear regression that the variance of the errors is constant is known as homoscedasticity. If the variance of the errors is not constant, it is called heteroscedasticity. If the errors are heteroscedastic, the estimated standard errors could be wrong.

White's test

To test for heteroscedasticity, we use White's test (1980, referred in Brooks (2008)). The test is based on obtaining the residuals from the regression in question. Then we estimate an auxiliary regression with the squared residuals as the dependent variable, and a constant, the original explanatory variables, the squared explanatory variables and their cross products as the independent variables.

$$\widehat{u_t}^2 = \alpha_0 + \alpha_1 x_{1t} + \alpha_2 x_{2t} + \alpha_3 x^2{}_{1t} + \alpha_4 x^2{}_{2t} + \alpha_5 x_{1t} x_{2t} + v_t$$

It is desirable to investigate whether the variance of the residuals varies systematically with any of the independent variables. The regression should include a constant term, even if the original regression did not. This is as a result of the fact that $\widehat{u_t}^2$ will always have a non-zero mean, even if $\widehat{u_t}$ has a zero mean (1980, referred in Brooks (2008)). R^2 is obtained and used in the test statistic, $T*R^2$, which is distributed as a $\chi^2(m)$. T is the number of observations and M is the number of independent variables in the auxiliary regression except for the constant. If the χ^2-test statistic is greater than the critical value, we reject the null hypothesis that the errors are homoscedastic.

ARCH-test (Engle's lagrange multiplier test)

The variance of the errors is unlikely to be constant over time in economic time series. Therefore, it makes sense to consider a model that does not assume the variance is constant, and which describes how the variance of the errors evolves. The volatility in economic time series often tends to cluster. In other words, large changes often follow large changes and small changes follow small changes. The current level of volatility tends to be positively correlated with its level during the immediately preceding periods (Brooks, 2008). This behaviour can be modelled using an ARCH model.

A test for determining whether 'ARCH-effects' are present in the residuals of an estimated model starts by saving the residuals from the regressed model. The residuals are then squared and regressed on q own lags to test for ARCH-effects of order q.

$$\hat{u}_t = \beta_0 + p_1 \hat{u}_{t-1} + p_2 \hat{u}_{t-2} + \ldots + p_r \hat{u}_{t-q} + v_1,$$

From this regression we obtain the R^2 and define the test statistic as $T*R^2$, which is distributed as a $\chi^2(q)$. T is the number of observations and R^2 is the coefficient of multiple correlation from the regression. The null –and alternative hypotheses are:

$$H_0: p_1 = 0, p_2 = 0 \dots p_r = 0$$

$$H_1: p_1 \neq 0, p_2 = 0 \dots p_r \neq 0$$

Potential difficulties with this test are the determination of the lag length and the possible violation of non-negative constraints.

7.3.6 Ramsey's reset test for misspecification

An assumption in the classic linear regression model is that the appropriate function form is linear. This means that the model is linear in the parameters. We can test whether this is upheld or not with the Ramsey's-reset test. The test uses higher order terms of the fitted values in an auxiliary regression. The regression can be written as:

$$y_t = \alpha_0 + \alpha_1 \hat{y}^2{}_t + \alpha_2 \hat{y}^3{}_t + \dots + \alpha_p \hat{y}^p{}_t + \sum \beta_i x_{it} + v_t$$

The regression contains the powers of the fitted values together with the explanatory variables. The higher orders can capture non-linear relationships, since they contain higher orders and cross products of the explanatory variables.

We obtain the R^2 for the test statistic, which is given by $T*R^2$. The test statistic is asymptotically distributed as $\chi^2(p-1)$. If the value of the test statistic is greater than the $\chi 2$ critical value, we reject the null hypothesis that the functional form was correct (Brooks, 2008).

7.4 Cointegration

Cointegration exists if two time series, which are characterized as I(1), can be combined as a stationary, I(0), linear expression. Many time series are non-stationary, but move together over time. The time series is connected to each other in a long-term perspective. If so, it exists a long-term equilibrium between the two time series. The time series can deviate from each other in the short run, but they will return in the long run. If the linear combination is non-stationary, the deviation from the long-run equilibrium will not be adjusted, and the deviation will be a permanent shock that deviates from equilibrium. The linear expression is called the cointegration vector. We can either estimate this vector or we can define it based on economic theory. A model with more than two variables can have several cointegration vectors.

If we want to analyse the relationship between two non-stationary time series, we can perform a regression with the first differenced variables. The result will be valid, but only in the short run. To find the long-term relationship, we have to see if the time series are cointegrated (Brooks, 2008). If the variables are non-stationary and not cointegrated, the statistical results may not hold and we can experience spurious results.

7.5 Error correction models (ECM)

A static model allows only variables from the same, period. In other words, changes in the independent variables at time *t* will immediately affect the dependent variable. An example of a static model:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + \varepsilon_t$$

A dynamic model will on the other hand take into account lags. A change in an independent variable today will not necessarily affect the dependent variable to a large extent today, but rather next period (t+1). In that case, we should include lagged values of the independent variable and possibly the dependent variable.

Models which include lags of both dependent –and independent variables, are called "Autoregressive distributed lag" models:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + \gamma_1 Y_{t-1} + \gamma_2 X_{t-1} + \gamma_3 Z_{t-1} + \varepsilon_t$$

If we subtract Y_{t-1}on both sides, as well as add and subtract $\beta_1 X_{t-1}$ and $\beta_2 Z_{t-1}$ (the total effect will only be minus Y_{t-1} on both sides), the total effect will result in the model expressed as below:

$$\Delta Y_t = \beta_0 + \beta_1 \Delta X_t + \beta_2 \Delta Z_t + (\gamma_1 - 1) Y_{t-1} + (\beta_1 + \gamma_2) X_{t-1} + (\beta_2 + \gamma_3) Z_{t-1} + \varepsilon_t$$

We can rewrite this to the following:

$$\Delta Y_t = \beta_0 + \beta_1 \Delta X_t + \beta_2 \Delta Z_t - (1 - \gamma_1) \left[Y_{t-1} - \frac{(\beta_1 + \gamma_2)}{(1 - \gamma_1)} X_{t-1} - \frac{(\beta_2 + \gamma_3)}{(1 - \gamma_1)} \gamma_3 Z_{t-1} \right] + \varepsilon_t$$

This model is called an error correction model. When transforming a static model to a dynamic model, possible autocorrelation will be reduced or disappear, as long as the autocorrelation is not caused by specification error.

An error correction model is a dynamic model, which studies the long-term equilibrium and how quickly deviations from equilibrium are corrected. The model contains both long-run - and short-run elements. An advantage with the error correction models is that the standard regression techniques are valid if cointegration exists, since the variables are stationary. There are several methods when estimating an error correction model. In this thesis we have used the Johansen test to estimate the error correction model. We will also present the much-used Engle-Granger two step method, because this is easier and more intuitive.

7.5.1 The two-step Engle-Granger method

Step 1 – the error correction term

We have to make sure the independent variables are integrated by first order, I(1). Then, we use the least squares method to estimate the cointegrating regression. The result may look like:

$$Y_t = \gamma X_t + \varepsilon_t$$

From the cointegrating regression, we save the residuals, $\varepsilon_t = Y_t - \gamma X_t$. We can now test if the residuals are stationary by performing an augmented Dickey-Fuller test of the residuals. If the residuals are stationary, the variables are cointegrated with the cointegration vector $[1, -\gamma]$, with γ being the cointegrating coefficient, and we can proceed to step 2. If the residuals are not stationary, there is no long-term equilibrium and we have to settle with a model only expressing the short-term relationships (Brooks, 2008).

Step 2 – Error correction model

In step 2, estimate an error correction model to find the equilibrium process. Use the residuals estimated in step 1 as one variable in the error correction model:

$$\Delta Y_t = \beta_1 \Delta X_t + \beta_2 (Y_{t-1} - \hat{\eta} X_{t-1}) + u_t = \beta_1 \Delta X_t + \beta_2 (\varepsilon_{t-1}) + u_t$$

Where the term in the brackets are the lagged residuals estimated in step 1, which are known as the error correction term (Brooks, 2008). In this case, the cointegrated vector is $[1, -\hat{\eta}]$, but this is just a linear transformation of the earlier cointegration vector. The interpretation of the model is that *Y* changes in the period between *t* and *t-1* caused by the changes in *X*. In addition, the model corrects any deviations from equilibrium in the previous period, *t-1*. At this stage, we can perform inferences based on the second stage regression concerning the parameters β_1 and β_2, since all the variables are now stationary. $\hat{\eta}$ defines the long-run relationship between *X* and *Y*, while β_1 describes the short-run relationship between changes in *X* and changes in *Y*. β_2 describes the speed of adjustment back to equilibrium, and its strict definition is that it measures the proportion of last period's equilibrium error that is corrected for (Brooks, 2008). B_2 will always have the value between 0 and -1. If the value is -1, we will experience an immediate return to equilibrium. If the value is 0, we will never return to equilibrium.

The Engle-Grangers method has several weaknesses. One of the problems is that the unit root –and cointegration tests lack power. Another problem is that the choice of dependent and independent variable, could affect the outcome through "*simultaneous equations bias*". This can occur if the causality between the variables runs both ways. We are forced to treat the variables asymmetrically by specifying one of the variables as the dependent and the other(s) as independent. In addition, the model can give skew estimates with few observations, but this problem disappears asymptotically. The final problem is the absence of the possibility to perform any hypothesis testing of the actual cointegrating relationship in step 1. This can result in misspecifications in the long-term equation in step 1 not being discovered and carried through to the cointegration test in step 2.

Despite the problems listed above, the model is popular in use. The easy use, the super consistency and the valid inference testing in step 2, are big advantages when using this model. Super consistency means that the estimated regression based on the cointegrated I(1) variables will reach the true value faster than stationary variables, when the residuals are stationary and the sample size increases.

7.5.2 Johansen test

The Johansen test is a procedure for testing cointegration of several I(1) time series. The test allows more than one cointegrating relationship, so it is more applicable than the two-step test. With g number of variables, up to g-1 number of cointegration relationships can exists. Johansen test can be used to test for the number of cointegrating relationships using vector error correction models. Suppose that a set of g variables are under consideration of being I(1) and which are thought may be cointegrated. A VAR with k lags containing these variables can be written as:

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots + \beta_k y_{t-k} + u_t$$

In order to use the Johansen test, we need to turn the VAR into an error correction model of the form:

$$\Delta y_t = \Pi y_{t-k} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} \ldots + \Gamma_{k-1} \Delta y_{t-k-1} + u_t$$

Where $\Pi = (\sum_{i=1}^{k} \beta_i) - I_g$ and $\Gamma = (\sum_{j=1}^{i} \beta_j) - I_g$.

This VAR contains g variables in first differenced form on the left hand side, and

k − 1 lags of the dependent variables (differences) on the RHS, each with a Γ coefficient matrix attached to it. The Johansen test centres on an examination of the Π matrix. Π can be interpreted as a long-run coefficient matrix, since in equilibrium, all the Δy_{t-i} will be zero, and setting the error terms, u_t, to their expected value of zero will leave $\Delta y_{t-k} = 0$ (Johansen, 1988).

The test for cointegration between the y's is calculated by looking at the rank of the Π matrix via its "eigenvalues". There are two test statistics for cointegration under the Johansen approach, which are formulated as:

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{g} \ln(1-\lambda_i)$$

and $$\lambda_{max}(r, r+1) = -T * \ln(1-\lambda_{r+1})$$

Where r is the number of cointegrating vectors under the null hypothesis and λ_i is the estimated value for the *i*th ordered eigenvalue from the Π matrix. Intuitively, the larger is λ_i, the more large and negative will $\ln(1-\lambda_i)$ be and hence the larger will be the test statistic. We reject the null hypothesis of there being r cointegrating vectors, if the test statistic is greater than the critical value. The alternative hypothesis is that there r+1 cointegrating vectors.

The first test has a null hypothesis of no cointegrating vectors, which is characterized by Π having zero rank. If this null is not rejected, it would be concluded that there are no cointegrating vectors and the testing would be completed. However, if $H_0 : r = 0$ is rejected, the null that there is one cointegrating vector (i.e. $H_0 : r = 1$) would be tested and so on. Thus the value of r is continually increased until the null is no longer rejected. Π cannot be of full rank (g) since this would correspond to the original y_t being stationary. If Π has zero rank, then by analogy to the univariate case, y_t depends only on Δy_{t-j} and not on y_{t-1}, so that there is no long-run relationship between the elements of y_{t-1}. Hence there is no cointegration. For $1 < \text{rank}(\Pi) < g$, there are r cointegrating vectors. Π is then defined as the product of two matrices, α and β′, of dimension (g × r) and (r × g) (Johansen, 1988).

For example, if we have one cointegrating vector(the β matrix), four variables(the y matrix) and the amount of each cointegrating vector entering the vector error correction model(α matrix, also known as adjustment parameters), matrix would be written as:

$$\Pi = \begin{pmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \\ \alpha_{14} \end{pmatrix} (\beta_{11} \quad \beta_{12} \quad \beta_{13} \quad \beta_{14}) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix}_{t-k} \quad \text{and rewritten as} \quad \Pi = \begin{pmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{13} \\ \alpha_{14} \end{pmatrix} (\beta_{11} y_1 + \beta_{12} y_2 + \beta_{13} y_3 + \beta_{14} y_4)_{t-k}$$

We can now write out the respective equations for each variable Δy_t. We first normalize the cointegration vector on the chosen dependent variable, resulting in the coefficient of that variable in the cointegrating vector is 1. Normalized on Δy_t:

$$\alpha_{11} \left(y_1 + \frac{\beta_{12}}{\beta_{11}} y_2 + \frac{\beta_{13}}{\beta_{11}} y_3 + \frac{\beta_{14}}{\beta_{11}} y_4 \right)_{t-k},$$

7.6 Hodrick-Prescott filter

The Hodrick-Precott filter (HP-filter) is a univariate mathematical model. An univariate model refers to a model of only one variable. The HP-filter uses only data from the actual time series to estimate the trend. The HP-filter can be expressed as $Y^t = \tau_t + c_t$. The time series Y^t consist of two components, the trend component denoted by τ, and the cycle component denoted by c. In this thesis we will only use this method to estimate the trend in GDP. We will therefore explain the method based on GDP to make it more intuitive. The method is based on finding the level of potential GDP, which minimizes the deviation between actual -and potential GDP. The model takes into consideration that the growth in potential GDP has a limited ability to fluctuate significantly. The HP-filter can be expressed as:

$$\sum_{t=1}^{T}(y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1}[(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2.$$

In this equation, y is the actual GDP and τ is the potential GDP. The first term in the equation is the squared deviation between actual –and potential GDP, which expresses the deviation from trend. The second term in the equation is the squared deviation of the growth in the trend component, which expresses the deviation in the trend. The HP-filter allows the trend to change over time, at the same time as it minimizes the deviation in the trend and the deviation between actual –and potential GDP. The multiplier λ represents the factor, which expresses how much we allow the potential growth to change. The value of λ can vary, the larger the value, the higher the penalty.

If the value of λ is very high, the first term in the equation becomes insignificant compared to second term. If chosen, we want to emphasize the minimization of the growth in the trend component. The trend will therefore be linear with a constant growth. This case is not very realistic, since we ignore possible shocks in the economy. On the other hand, if we set λ=0 the second term in the equation equals 0. If chosen, we only minimize the deviation between the potential –and actual GDP. This means that $Y^t = \tau_t$ and the trend coincides with the actual values and the output gap will always be 0. This is also unrealistic, since we ignore the existence of business cycles.

Which value we chose to use depends on what kind of data used and the interval of time between the measurements. Yearly –and quarterly data for the same data series have different optimal values of λ. There are several advantages and disadvantages with the HP-filter (Klovland, Helliesen and Kvåle, 2012). The model is easily applicable and intuitive, as well as we only need one time series for estimation. The disadvantages with the model are the sensitivity of the chosen value for λ, which can make the results differ significantly. The larger the value of λ, the bigger the end-point problem (Bjørnland, Brubakk and Jore, 2004). There are several suggestions of the value of λ. Kydland and Prescott (1990) suggested to use λ=1600 on quarterly data from USA. This resulted in a trend that was reasonable taking other observations into account, which has led to that λ=1600 is the most commonly used in literature for quarterly data. Statistics of Norway has estimated that a λ-value of 40000 is more descriptive for Norwegian GDP data (Johansen and Eika, 2000).

Another problem is that the data in the start –and the beginning of the series are much more affected by the actual data than the rest of the series. This is because the model uses data from both the most recent period (y_{t+1}-y_t) after, and before (y_t-y_{t-1}), period t. In the end-points, the model can just use one of the above. This results in the first –and last trend estimate to be unreliable.

The length of the business cycles will limit how well the HP-filter estimates the trend. The model weights positive and negative deviations equally, which implies that economic upturns and downturns are equally long on average. This does not necessarily correspond with empirical studies (Romer, 1999).

It is also important to remember that the HP-filter is a pure mechanical method to estimate trend and it is not based on economic theory.

We will now start our statistical analysis of the Norwegian housing market, based on the statistical methods presented in this chapter.

8. Building a model for the development in Norwegian housing prices

We want to create a new model that could help explaining the development in the Norwegian housing market. We have used economic theory, evidence from previous studies and previous arguments in the thesis when choosing the components in the model.

When performing statistical analysis of economic time series, we always have to ensure that the parameters correspond to economic theory. Our general knowledge and perception of market forces and how different variables affect each other, is very important in order to generate a reliable and applicable model. Sometimes it might be correct to include variables that are statistically negligible because they are supported by economic theory.

8.1 Presentation -and expectations of the variables

We have emphasized economic theory when including and discarding variables in our thesis. Therefore, it is important to identify the economic theory and our own expectations regarding the variables before we start modelling. When mapping our expectations regarding the variables' stationarity, we will ask ourselves the following question: "How will the variable develop in the long term when a exogenous shock occurs?".

We believe the real interest rate after taxes, unemployment rate and housing stock are variables affecting the housing prices negatively. These variables create uncertainty concerning future disposable income and living expenses. On the other hand, we believe that the variables income and consumers expectations will have a positive effect on housing prices.

We will include the following variables in the further testing and modelling:

1. *The real housing prices*. These are based on a hedonic housing price index measuring average housing prices in Norway. The index is calculated on the basis of sales data in the second hand market. Statistics Norway officially started publishing housing price data in 1992. Prior to 1992 an unofficial index, based on similar sources and compiled at Statistics Norway, is used. A shock in the

housing prices can occur if there are changes in tax-policies or regulations affecting the housing market. If these changes are permanent and affecting the prices negatively, we believe that the prices will stabilize at a lower level than before the change. Therefore, we expect the housing prices to be a non-stationary variable.

2. *The households' total disposable income adjusted for capital income*. We have chosen to subtract capital income, since this can be affected by changes in tax policies. An example of this is the tax reform of 2006, which resulted in an abnormal capital income among households in 2005.
   The data is gathered from Statistics Norway. Higher income will result in increased purchasing power and borrowing capacity among households, which increases the amount of money disposable to spend on consuming goods and investments. If the nominal wage-growth is higher than expected, it can be regarded as a shock in income. We find it hard to believe that the wage-growth will fall, but rather that the workers expectations for future wages will adapt to the new wage-level. Since the effect of a shock in income will not die out, we believe this variable to be non-stationary.

3. *The households' total amount of outstanding gross debt.* The data is gathered from Statistics Norway. When discussing how the households' gross debt affects the housing prices, we are not completely certain. We believe that increased housing prices could result in higher gross debt among households, since they can borrow more with the dwelling as collateral. On the other hand, we do not find it likely that the housing prices increase due to an increase in gross debt. If the gross debt changes, we suggest four possible causes: a change in income, housing prices, bank's lending policies or interest rate. We see the changes in debt more as a result of changes in other variables, than debt being an independently exogenous variable affecting the housing prices directly. If this is the case, the model can experience simultaneous inference. This can cause the residuals to be correlated with one or more of the independent variables. This violates one of the standard assumptions for OLS. Three of the four causes are individual variables already discussed. To see if the fourth option, banks'

lending policies, have impact on the housing prices, we need to see if the credit supply is affected by the banks' profit or government regulations. If we do not find households' gross debt to be a significant variable explaining the housing prices, it will support the theory that the lending was not limited by the banks' profit or government regulations during our estimation period. The households' debt is included as a significant variable with positive effect in Norwegian housing models from the 1980s –and 90s (Eitrheim (1993), referred in Jacobsen and Naug (2004:2))(Boug and Dyvi, 2002). Anundsen and Jansen (2011) find evidence that it exists a self-reinforcing relationship between debt and housing prices. The results establish a two-way interaction in the long run, meaning that higher housing prices lead to a credit expansion, which in turn puts an upward pressure on prices. We choose to include the debt variable in further analysis and modelling, but we are not certain of its effect on housing prices. We expect the variable to be non-stationary, because we do not find it likely that the debt will decrease after a positive shock. For example, a positive shock can be caused by changes of government –and banks' lending policies.

4. *Real housing stock measured in fixed prices*. This measures the total stock of housing in Norway and it is calculated according to the perpetual inventory method[5]. The data is gathered from Statistics Norway. We believe the development in initiated dwellings and building costs are integrated in the development of the total housing stock. We will therefore not include these in the further analysis, but rather use housing stock only. An increased housing stock will increase the supply of housing and therefore reduce the price pressure in the market. We believe that the housing stock will not have a short-term effect on the housing prices. This is because economic theory suggests that the short-term housing supply is fixed due to limited capacity in the construction industry. A shock in the housing stock can be caused by drastic changes in climate and housing requirements, making dwellings uninhabitable. The change in housing stock will be permanent, but it is likely to believe that the

[5] System of inventory control in which the number of units are continuously recorded to provide a running balance.

negative shock will be followed by a positive shock due to new initiated dwellings. Since the shocks accumulate over time and results in continuously growth in housing stock, we believe the variable is non-stationary.

5. *Real interest rate after tax calculated by adjusting for inflation based on the consumer price index*. The nominal interest rate used is an average of the rates paid by households on loans in private financial institutions. The nominal interest rate is adjusted by the marginal tax rate and deflated by CPI. The data is gathered from Statistics Norway. We chose to use CPI as our deflator. We calculated the real interest rate based on all three deflators. The real interest rate deflated by CPI did not fluctuate as much as the others and we believe this one is better to use further in the analysis. This is also the deflator used by the Norwegian central bank. A shock in the interest rate will probably have great impact on the economy. The key policy rate is a powerful tool that central banks use to affect the interest rates in the market, and hereby correct unwanted market development such as inflation and overheating in the economy. When the economy returns to normal circumstances, the interest rate level will eventually return to the "normal level". On this basis, we expect the real interest rate to be stationary. We have to keep in mind that we are examining a limited time period, and the time series can be non-stationary even though economic theory indicates such variables to be stationary.

6. *The seasonally adjusted unemployment rate*. The data is gathered from Statistics Norway. Changes in regulations can result in a shock in the labour market. The theory of hysteresis suggests that an abnormal high or low unemployment rate can settle, and not automatically return to pre-shock level (Ministry of Finance, 2000). This may be caused by the increased difficultness of finding a new job after being unemployed for a long period of time. Long-term unemployment can result in changes in personal productivity, motivation and abilities. A study by Røed (1996) finds evidence of the theory of hysteresis in most European countries, including Norway, when analysing the stationarity of the unemployment. However, these tests do not consider any exogenous effects that can result in a permanent change in labour market equilibrium, such as

structural reforms. Taking this theory into account, we believe that shocks in the labour market can have long-term effects on the unemployment rate. Therefore, we expect the variable to be non-stationary.

7. *The households' expectations and economic confidence*. The expectations variable is constructed by TNS Gallup, and can be seen as a consumer confidence indicator. It is based on a survey, where average score can range between −100 and 100. The indicator measures households' expectations concerning the state of the economy and the development in their personal economy. People's expectations towards personal –and the nation's economy fluctuate with the business cycles. If the economy is in a boom and then suddenly experiences a shock, for example terrorist attacks or natural disasters, people's expectations can suffer a drastic fall. We believe that with a longer time perspective, these negative expectations will not sustain. People are often optimistic and tend to forget negative episodes quickly. Therefore, we expect the expectations variable to be stationary.

Regarding the development in gross domestic product, we have decided not to include this as an independent variable. This is because of the many problems regarding the use of GDP mentioned earlier in the thesis, and the possible weakened strength and validity of our final model. Although, we will use GDP later in the thesis, when analysing the business cycles' effect on housing prices.

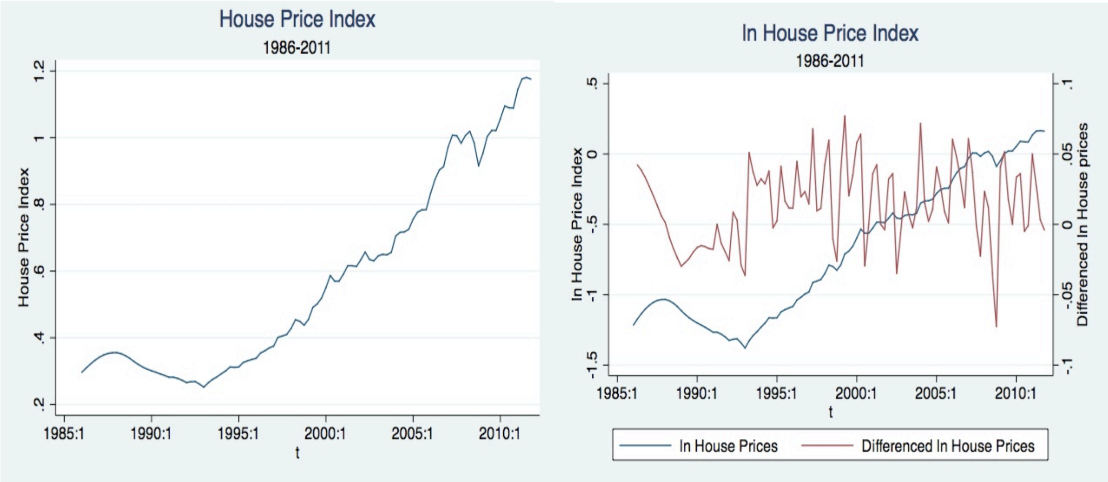
8.2 Stationarity analysis

In the development of a new model, we need to see how the housing prices develops and try to find reasonable arguments for the influence of different variables. We want to find out whether we can assume the data to be stationary or not. By plotting the data, it might be possible to reveal possible trends, seasonality or other peculiarities.

8.2.1 Housing prices

If we start by looking at the housing prices, we have already talked about the huge increase the last 20 years. When we want to explain the development by the impact of other variables, it might also be valuable to take a look at the housing prices once more. In figure 18 the price index from SSB (2012:3) is plotted from 1986-2011. We see that the growth has been substantial in this period, but it does not seem to be linearly. This makes sense, since the housing prices are affected by several factors in the market. We cannot assume a stable absolute growth in a variable affected by so many relations and factors as the housing prices. Over a long period of time, we might assume a more stable development, due to inflation and long-term changes in wealth. We would not assume the price to increase with the exact same value every year, but rather with the same average percentage, like the 2,5% inflation target set by the Norwegian central bank. This would suggest an exponential growth in the housing prices, which might also seem reasonable when evaluating figure 18 graphically. The variable plotted on a logarithmic scale in more linear. We chose to use this in further modelling and analysis.

Figure 18. LHS Norwegian house price index. RHS Norwegian house price index, logarithmic and differenced (SSB)



We also have to test whether or not the variables seem stationary. As mentioned earlier, a variable is weakly stationary if the mean, variance and autocorrelation are constant. All test results can be found in the appendix.

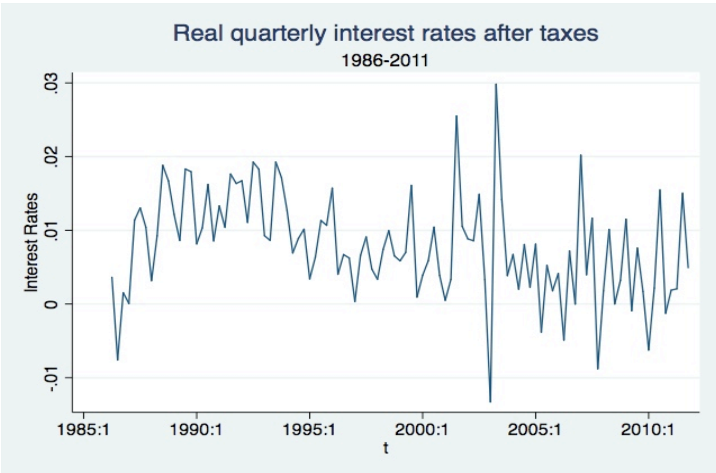
In figure 18, we can see the plot of the housing prices together with the plot of the first differenced housing prices. When using the ADF-test we can choose to include both trend and constant, only constant or neither. We also need to specify the number of lagged differences we want to include in the test. We can see that housing prices seems to have a pretty clear trend, and a constant. The number of lags might be decided in several ways, but we have used the Schwarz Bayesian Information Criterion (SBIC) to decide the optimal lag length. When testing with lags, we have to make sure that the last lagged difference has a coefficient that is significantly different from zero. If not, this extra variable results in less degrees of freedom, which decreases the power of the test. We follow the testing strategy suggested by Elder and Kennedy (2001) while using a 5% significant level. SBIC suggested an optimal lag length of 5 lags. We expected this variable to be non-stationary, and the test confirms our intuition by failing to reject the null hypothesis of the variable being non-stationary.

The next step is to difference the variable, and test whether the variable is I(1). We tested with a constant, but no trend since this seemed reasonable according to figure 19. The constant turned out to be insignificant, so we tested with neither trend nor constant and found the first differenced of the housing prices to be stationary. We therefore conclude with housing prices being an I(1)-variable.

8.2.2 Interest rates

Examining figure 19, the interest rates seem to be stationary, but there might be a small declining trend and a possible drift. Since we are using quarterly data, the real interest rate might fluctuate more than annual data.

Figure 19. Real quarterly interest rates after taxes (SSB).

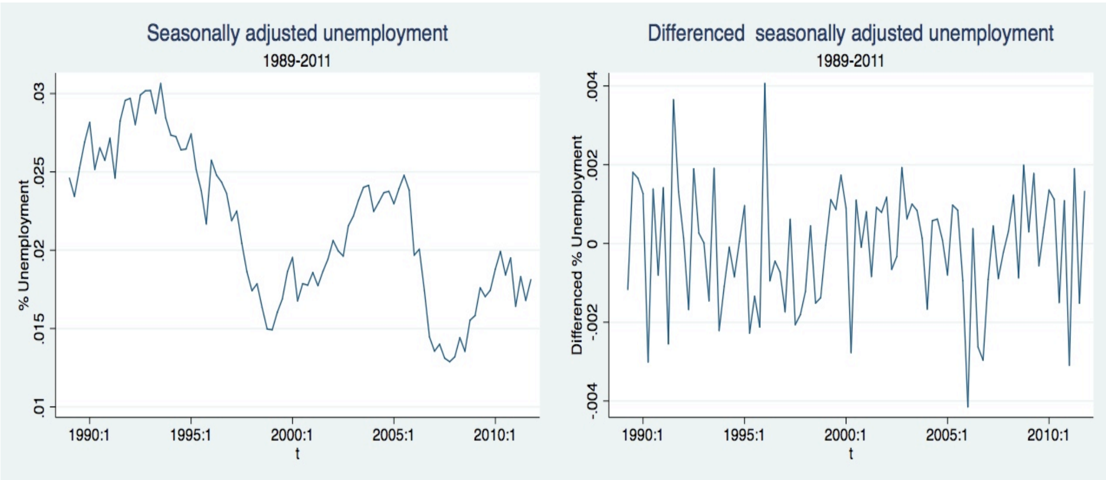


SBIC suggested 1 lag. We expected this variable to be stationary, and the test confirms our intuition by rejecting the null hypothesis. We therefore conclude that the real interest rate seems to be an I(0)-variable.

8.2.3 Unemployment

It seems like the unemployment is showing a declining trend according to figure 20, so we include a trend and constant when running the ADF test.

Figure 20. LHS Seasonally adjusted unemployment rate (SSB). RHS Differenced seasonally adjusted unemployment rate



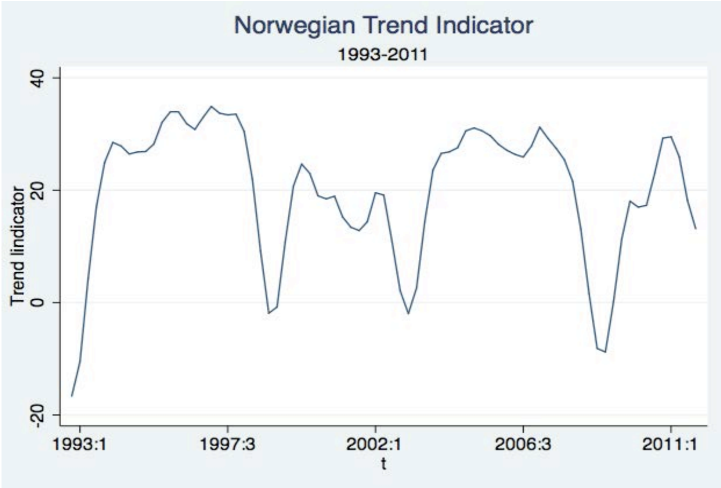
SBIC suggested 1 lag. We expected this variable to be non-stationary, and the test confirms our intuition by failing to reject the null hypothesis.

We then need to test the differenced value, to see if the variable is I(1). In figure 21, we have plotted the differenced variable. It looks stationary with no trend and a possible drift. We ran the ADF test with 0 lags and we could reject the null hypothesis. We therefore conclude that the unemployment seems to be an I(1) variable.

8.2.4 Trend indicator

From the plot in figure 21, we can see that the variable seem to have no trend and a possible drift, so we include a constant in the ADF test with 0 lags as suggested by SBIC. We expected this variable to be stationary, and the test confirms our intuition by rejecting the null hypothesis.

Figure 21. Norwegian trend indicator. 1993-2011. TNS Gallup

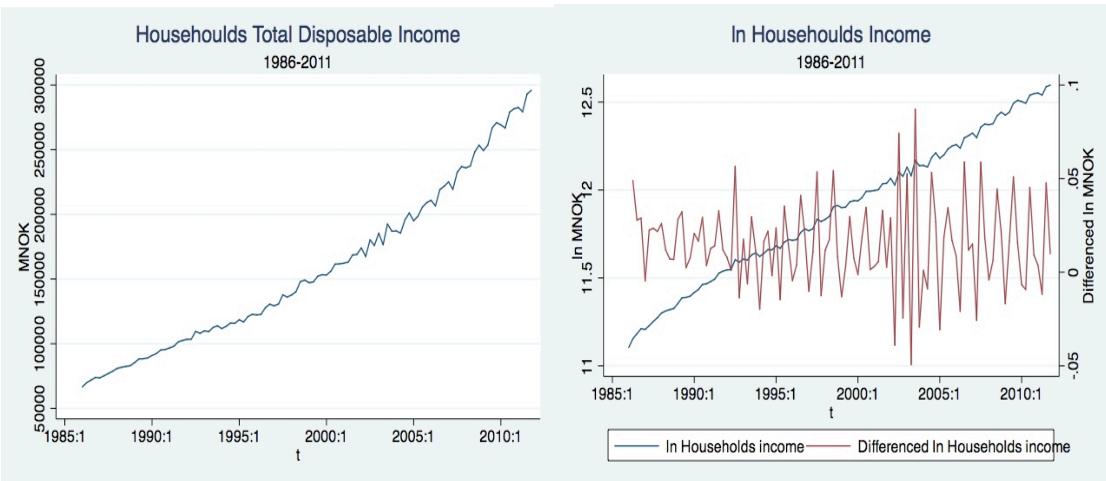


We therefore conclude that the trend indicator seems to be an I(0) variable.

8.2.5 Income

We can see from figure 22 that the households' disposable income has increased steadily, and somewhat exponentially, especially since 2000.

Figure 22. LHS Household's total disposable income (SSB). RHS Household's total disposable income, logarithmic and differenced



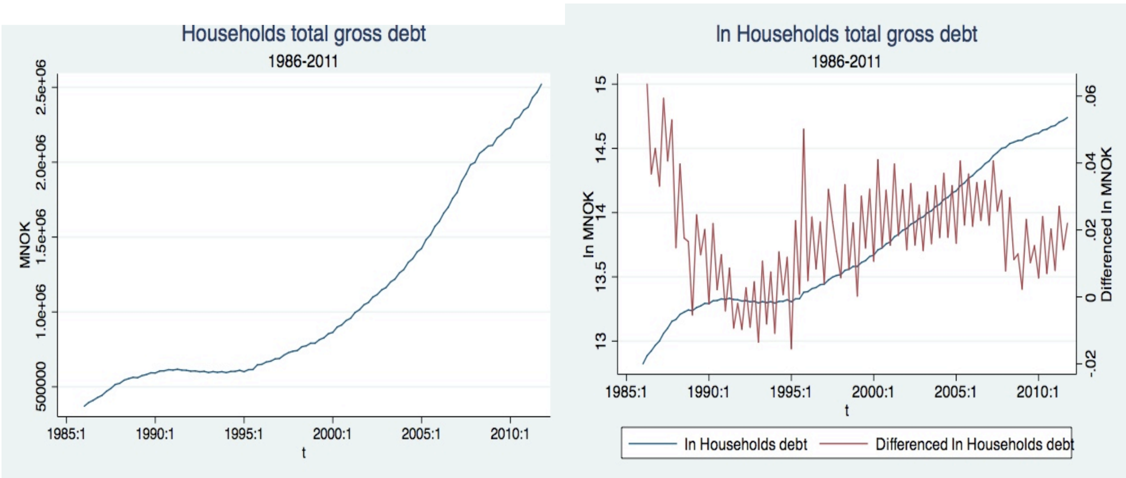
From the series plotted on a logarithmic scale we can see that the variable expressed in a logarithmic term gives a better fit to a straight line and will therefore be used in further analysis and modelling. The series clearly shows a deterministic trend and we chose to run the ADF test with both constant and trend, with 4 lags. We expected this variable to be non-stationary, and the test confirms our intuition by failing to reject the null hypothesis.

We then need to test the differenced values. With 5 lags, we rejected the null hypothesis. We conclude that the households' disposable income seems to be an I(1) variable.

8.2.6 Household's gross debt

We can see by the plot in figure 23 that the Norwegian households' gross debt has had a somewhat exponential growth since the 1990s.

Figure 23. LHS Household's total gross debt. RHS Household's total gross debt, logarithmic and differenced (SSB)



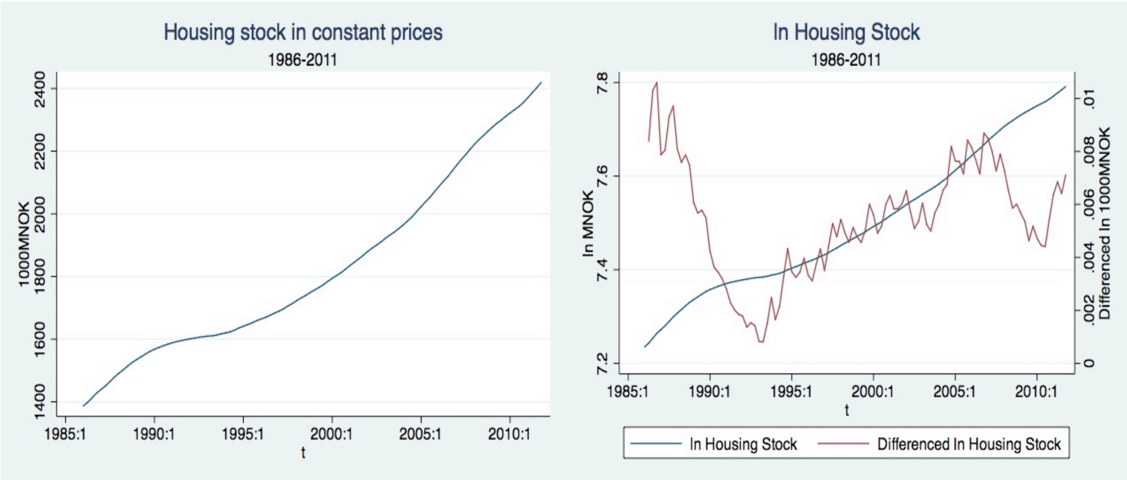
We can see that the debt expressed in a logarithmic term gives a better fit to a straight line and will therefore be used in further analysis and modelling. The series shows a clear deterministic trend and possible drift, so the ADF test is run with a constant, trend and 5 lags. We expected this variable to be non-stationary, and the test confirms our intuition by failing to reject the null hypothesis.

We then need to test the differenced values. With a constant and 4 lags, we rejected the null hypothesis and conclude that the variable seems to be an I(1) variable.

8.2.7 Housing stock

The housing stock in fixed prices has been increasing steadily throughout the period. There are indications of an exponential growth from 1990.

Figure 24. LHS Housing stock in constant prices (SSB). RHS Housing stock in constant prices, logarithmic and differenced



By looking at the series plotted on a logarithmic scale, we can see that this fits a straight line better than the non-log series. Therefore, we choose to use this variable on a logarithmic scale in the further analysis. When looking at the logarithmic plot, we identify a clear trend. We run the ADF test with trend, constant and 4 lags. We expected this variable to be non-stationary, and the test confirms our intuition by failing to reject the null hypothesis.

We then need to test the differenced values. When testing with no trend or constant, and 4 lags, we could reject the null hypothesis. We conclude that the housing stock seems to be an I(1)-variable.

8.3 Determining the phases of the business cycle

By decomposing GDP into trend and cyclical values, we will try to use the business cycles in a model explaining the housing prices. We will start with determining the different states in the business cycles over the last 35 years. The period is chosen on the basis of the quarterly data available from Statistics Norway. We wanted a bigger dataset than just the estimation period from 1986-2011, since this will provide us with more observations when estimating the trend.

We started with collecting the quarterly GDP mainland data for Norway from Statistics Norway in fixed prices. The quarterly data available was in the period from 1978 until the second quarter of 2012. We then used the software Stata, to seasonally adjust the data. This is necessary to describe the trend without the seasonal fluctuations. We used the methoddescribed earlier in the thesis, which gave us the seasonal weights in figure 25.

Figure 25. Seasonal weights

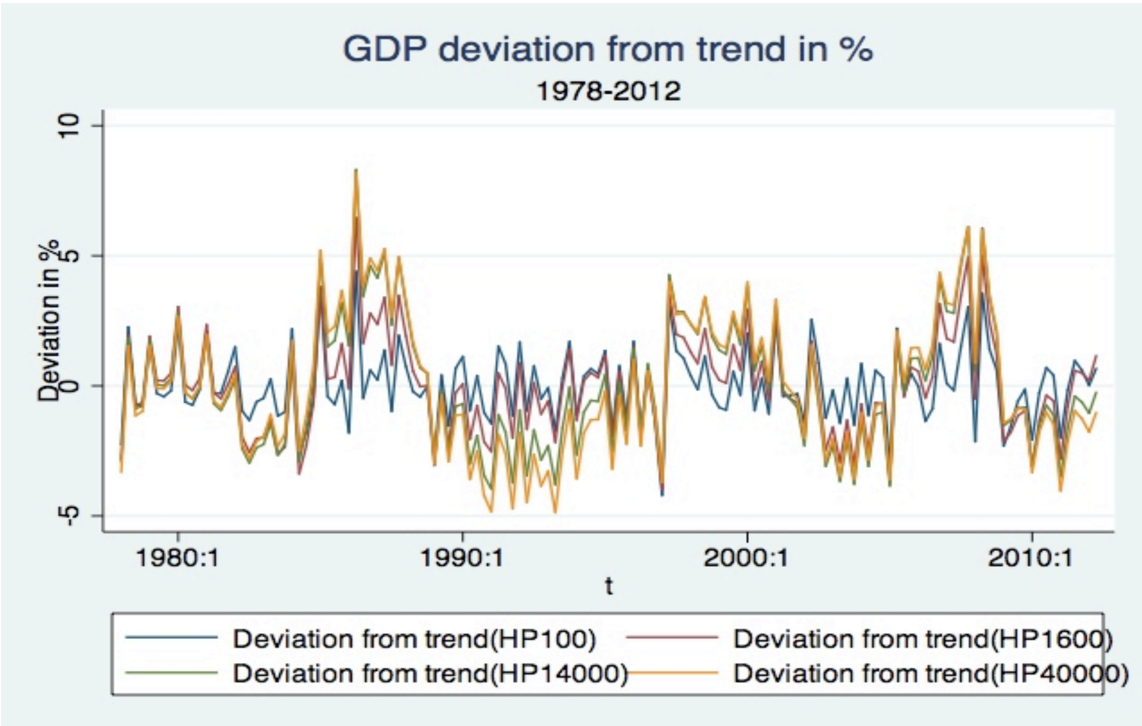
Quarter	**Seasonal weight**
1	0,975320
2	0,983816
3	1,004159
4	1,036695
Total	**4**

The seasonal weights represent the quarterly adjustment factor. We divide the real quarterly GDP by the adjustment factor to seasonally adjust the data. As we can see, GDP is usually lower in the first six months, so we increase the first two quarterly data slightly while we reduce the last two. This resulted in a smoother estimate of the development in GDP.

We assume that the trend in GDP is deterministic. If we then de-trend the series, we can more easily see the deviation from trend. To do this we need to estimate the trend. This can be done in many ways, like linear - or exponential regression, but we have chosen to use the Hodrick-Prescott filter. This is one of the most common ways of de-trending when working with business cycles. The problem is to know which value of Lambda to apply. We therefore tried four commonly used values of Lambda to de-trend the logarithmic values of quarterly seasonally adjusted GDP-data. We used historic numbers from other papers and articles to try to determine which of the trends was the most accurate describing the Norwegian business cycles. Next, we calculated the trend development for the different values of Lambda, and estimated the percentage deviation from the seasonally adjusted real values of GDP. The business cycles were then estimated as the deviation from the trend. The different results where plotted in figure 26 to try spotting differences. It is important to keep in mind that there is a lot of uncertainty connected to the method, and we have to be critical when interpreting the results.

The option with using lambda value of 40 000 fitted the data best, which is also concluded in earlier works (Johansen and Eika, 2000). For example, in the economic downturn at the beginning of the 1990s the common opinion is that the trough in the business cycles was early in 1992. The negative gap in production calculated with λ=1600 reaches the trough already in 1989 while the gap in production calculated with λ=40000 reaches the trough in 1991/92.

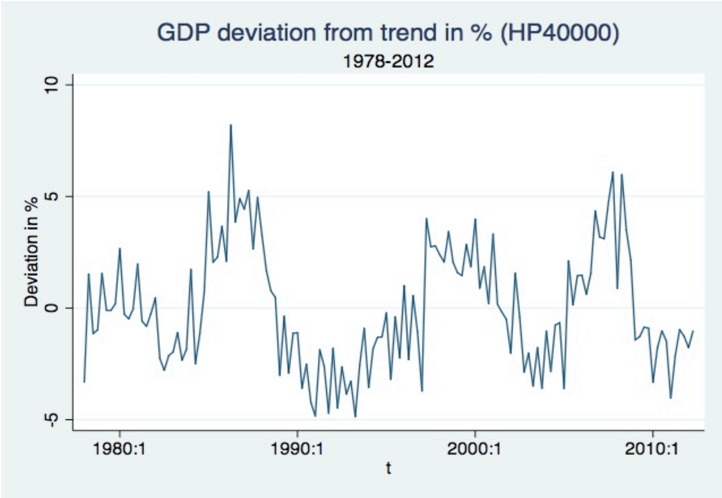
Figure 26. GDP deviation from trend with different HP values.



40 000 as the value of Lambda matches our knowledge and views of the Norwegian business cycles the best, also on the other well known troughs and peaks. Therefore we decided to apply this parameter value.

There are still a lot of fluctuations in the series, and it might be difficult to see the changes in development and determine peaks and troughs in the cycles. To smooth the development and make it easier to work with, we calculated a moving average of ±2 quarters for every quarter in the seasonally adjusted data. We then calculated the percentage deviation from the trend using the HP-filter with Lambda = 40 000. This is illustrated in figure 27.

Figure 27. GDP deviation from trend, HP=40000



In the table beneath, we have tried to classify the business cycles in the period from 1978 to 2012q2.

Figure 28. The quarters categorized by the different phases within a business cycle

From	Until	Overheating	Cooling	Setback	Recovery
1978:3	1980:3	X			
1980:4	1981:3		X		
1981:4	1983:1			X	
1983:2	1984:3				X
1984:4	1986:4	X			
1987:1	1988:3		X		
1988:4	1992:4			X	
1993:1	1996:4				X
1997:1	1997:4	X			
1998:1	2001:3		X		
2001:4	2003:2			X	
2003:3	2005:2				X
2005:3	2007:2	X			
2007:3	2009:1		X		
2009:2	2010:3			X	
2010:4	-				X

We have evaluated the data as a growth cycle. This means that when there is a peak in figure 27, there is a peak in a business cycle seen as a growth cycle. This is characterized as the point when the cycle is deviating the most from the trend, which is the same as when the derived of the cycle is equal to the derived of the trend. These dates match pretty well with our historical presentation in the beginning of the thesis.

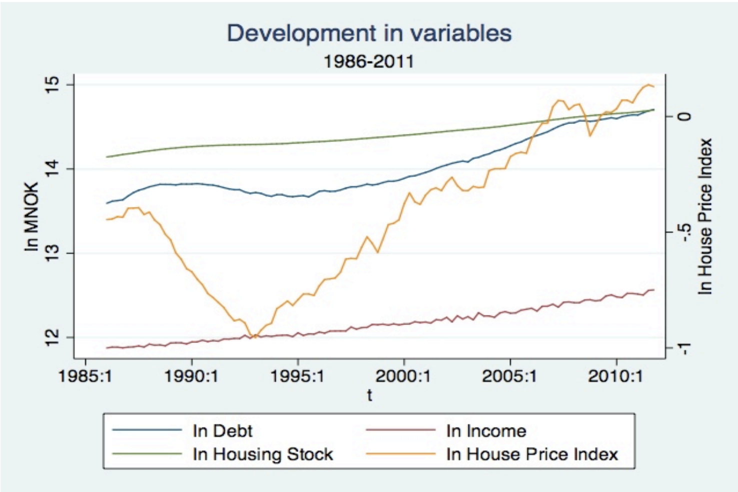
We have now classified the last 35 years in different phases of a business cycle. In the next section, we will use this information further, when building a model describing the Norwegian housing market.

8.4 Cointegrated variables

The general rule when using ordinary least squares method is to only apply stationary variables. This is because using non-stationary variables might lead to spurious results. The exception is if any of the I(1)-variables are cointegrated. Two cointegrated variables have a common stochastic trend, and will never diverge too much from each other. We use the Johansen test in Stata to find the possible number of cointegrated equations between the variables in question (Johansen (1988), referred in Brooks (2008)). The tests shows that the following I(1)-variables could be cointegrated:

- Housing prices
- Housing stock
- Income
- Debt

Figure 29. Development in the variables on a logarithmic scale



These four variables are illustrated in the logarithmic plot in figure 29.

When testing the cointegrating relationships, we use the Johansen test. To decide the appropriate number of lags and whether or not to include a constant term, we use information criterions and economic theory. If a constant is included in the cointegration vector, it assumes a drift in the relationship between the variables. This is because the constant is included in explaining the differenced dependent variable. The change from one period to the next will therefore always include a constant term. We have noticed that neither of the two housing price models presented earlier in the thesis has included any constants in their cointegrating relationships. This subject is not well covered in the literature and empirical research, but we will try to use economic reasoning to evaluate our cointegrating variables (Ahking, 2001).

It is important that the cointegrating variables have a plausible relationship. We will therefore take a closer look at the most interesting combinations.

Housing prices and debt are one of those relationships. Most houses in Norway are debt financed, and a large share of the households' total debt is mortgages. With increasing housing prices, the possibility of refinancing the debt has also increased. In this way, the mortgage can be increased to purchase other consumable goods. The government is trying to regulate the credit market to avoid the debt increasing along with the housing prices. The 15% equity rule is a good example. Still we find it reasonable to believe that these two variables will have a pretty stable relationship, and that they might be used as cointegrating variables.

We find it reasonable to believe that the housing prices and debt affect each other in a two-way relationship. If the housing prices increase, the demand for debt will also increase. But if the price on debt, the interest rate, or the availability of debt changes, it might also change the housing prices. It is difficult to know whether or not to include a constant in the cointegrating relationship. We have decided not to, since we are uncertain of the development between the debt and the housing prices, and cannot claim that there is a fixed element in their development. The information criteria suggest 5 lags when we test without including a constant. When g lags are included in the VAR (Vector Autoregressive model), this gives $g-1$ differenced lags in the VECM (Vector Error Correction Model), and it seems natural with 5 lags initially since we are operating with quarterly data.

The relationship between housing prices and income has been closely evaluated in previous articles, and is included in both models presented earlier in this thesis. Higher income will increase the amount of money available for housing services, and probably increase prices. Likewise, if income is reduced the demand for housing services will probably decrease, and the prices fall. Hence, it seems reasonable that these two variables might be cointegrated. When looking at the development between the housing prices and income graphically in figure 29, it seems difficult to say something certain about the relationship. We believe they are cointegrated due to economic reasons, but we will not include a constant in the cointegrating equation since we do not have any clear economic reasons to do so. We also assume that housing prices will depend on income and not the other way around. The information criteria suggest 5 lags.

Income and housing stock seems to have a pretty stable relationship according to figure 29. The time horizon is important because of the characteristics of the housing stock, since it is seen as given in the short-run. The income is also difficult to affect in the short-run, because of unions and wage negotiations. In the long run they are both more dynamic. It is still reasonable to assume a stable relationship between the two variables, since income logically will have a major impact on the housing stock. We do not see a clear economic reason to include a constant in the relationship between income and housing stock, so we will therefore not include one. We will estimate the housing stock as a function of the income, because this is the most reasonable relationship. The information criteria suggest 5 lags.

We use the Johansen test to find the cointegration vector. It also shows the speed of adjustment, the coefficient estimate of the error correction equation, and tests if this factor is significantly different from zero. The test output can be found in the appendix. All factors proved to be significant, and we found the following equations:

- $ce\_ph\_d = ph + 0{,}0932852d$
- $ce\_ph\_yd = ph + 0{,}1483129yd$
- $ce\_h\_yd = h - 1{,}206192yd$

where

- $ce\_x_1\_x_2$ = Correction error equation for integrated variables x_1 and x_2
- ph = ln housing prices deflated by the consumption deflator

- d = ln households gross debt deflated by the consumption deflator
- h = ln Housing stock in constant prices

We see that the cointegration equation including housing prices and income seems to contradict economic reasoning. No matter the sign of the coefficient in front of the vector, equal signs of ph and yd means that increasing income will contribute to decreasing housing prices. We have seen that this relationship has been set to "$1 - 1$", or "$\mathrm{ph} - \mathrm{yd}$" in the MODAG model for housing prices presented earlier in the thesis. By doing so, they assume homogeneity between the variables and restricts the effect of income to positive. We will although wait until we have the total affect from the cointegrating equations before we make any decisions of whether to drop this equation or not in our model. The income is also included in a cointegrating relationship with the housing stock, and the opposite effect on the housing prices from these two follows economic reasoning. The housing prices and the debt in the first cointegration should also have opposite signs to follow economic reasoning. Housing prices are although included in two different cointegrations, so the total effect is still unclear and dependent of the coefficients in the final model. When we have more than one cointegrating equation including the same variables, the total effect from the variables might be different from what it seems like in one of the cointegrating equations alone. We will therefore use these equations in our model as a part of the initial error correction model, and then test the total affects before deciding on including them or not.

In this chapter we have analysed and processed the data, in order to make sure the variables are suited for further modelling.

9. Estimating a model for the housing market

We now have the explanatory variables we want to use and the error correction equations. The stationarity tests indicates which other variables we can include, and on which form. Since the dependent variable housing prices is an I(1)-variable, it must be on differenced form. The regression will indicate the effects of several variables on the change in the housing prices. We must always remember that the coefficients are estimated by mathematical models and not economic reasoning. We therefore have to be careful when interpreting the results, and make sure the relationships are theoretically grounded. The easiest way to spot if the numbers are reasonable is to check if the sign of the coefficients match our expectations.

An important decision is whether or not to include a constant in the regression. A constant in the equation of a differenced variable will mean that we believe the variable has a constant drift. If all explanatory variables remain unchanged, the housing prices will still change due to the constant. Looking at the housing prices in our data, it might seem like we have a positive trend. The prices have risen more than the variables in this period, but it is difficult to find out if it would grow without any change in the other variables. There could also exist variables that we are not including, or effects our variables do not pick up, making the housing prices grow even if the included variables remain unchanged. We will include quarterly dummy variables to account for seasonal fluctuations. If we exclude a constant while including seasonal dummies, the dummies might end up including non-seasonal effects. We will therefore include a constant in our model.

When building a model like this, there will always be tradeoffs between simplicity and explanatory power. This is especially important when deciding how many lags of the variables we want to include. Too many explanatory variables might also tend to weaken the power of one another.

The differenced housing prices can be expressed as $\Delta ph_t = ph_t - ph_{t-1}$. We believe that it exists inertia in the housing market. This means the prices uses some time to adapt to changes in the factors affecting the market. This is the reason why we use the error correction model to include the deviations from earlier periods. Since the dependent variable is dependent of the previous period, we use the variables on lagged form. We will also include more lags of the explanatory variable. Dummy variables are included to try to capture the effects of seasonal variations. The constant will include the affect of the fourth quarter, and we will maximum need three quarterly dummies. We will also include a dummy variable indicating if the GDP is over the trend or not, stating if there is a boom or a recession.

If we were to start with a balanced model, which means including the same number of lags of every variable in the short run dynamic, the model would be over-parameterized. We therefore have to base our modeling strategy on reducing the balanced model to a more specific model with reasonable economic and econometric qualities (Jacobsen and Naug, 2004:2). When choosing how many variables to include, it is a tradeoff between inaccurate and skewed estimates. If we include many variables and lags we will most likely end up with a model that is not too exposed for skewness, which means the probability of omitted variables is reduced. On the other hand, including many variables will increase degrees of freedom, which might create inaccurate estimates.

$$\begin{aligned}\Delta ph_t = {} & \sum_{i=1}^{4}\beta_{1,i}\Delta ph_{t-i} + \sum_{i=1}^{4}\beta_{2,i}\Delta h_{t-i} + \sum_{i=1}^{4}\beta_{3,i}\Delta d_{t-i} + \sum_{i=1}^{4}\beta_{4,i}\Delta yd_{t-i} + \beta_5 ce_{ph\_yd,t-1} \\ & + \beta_6 ce_{h\_yd,t-1} + \beta_7 RRT_{t-1} + \beta_8 Seas_t + \beta_9 Seas_{t-1} + \beta_{10} Seas_{t-2} \\ & + \beta_{10} DummyGDP + constant\end{aligned}$$

If we start with a model including all variables with the appropriate number of lags and then remove variables that are insignificant[6] or in conflict with economic theory, the order in which we remove variables might have a substantial impact on the outcome. The same problem will occur if we start with a few variables that we know we want to include, and then add more variables as we go. To ensure that we do not wrongly omit any variables or end up with spurious results (Freedman, 1983), we have estimated several different models including only a few of the variables. In the evaluation of each model, we emphasize

[6] Coefficient significantly different from zero, hereby p-value < 0.05, are significant and can be used in the model

economic theory and simplicity, rather than just looking at predictability and significance. Then, we simplified these models by restrictions. Some of the restrictions applied were not rejected and eased the interpretation of the dynamic between the variables. By analyzing the dynamics between the variables in this way, we could enhance our certainty when reducing the number of variables.

The Norwegian trend indicator was supposed to capture the influence from the consumers' expectations of the housing prices. The variable turned out to be significant, but often with a coefficient that was negative or extremely close to zero. A negative coefficient would imply that positive expectations would decrease the housing prices, which does not seem correct according to economic theory. The coefficient close to zero would not have made any impact no matter what, so we excluded this variable.

The differenced debt never turned out to be significant, nor did the coefficient for the cointegrated vector including the debt. This can be the result of the theory that the lending was not limited by the banks' profit or government regulations during our estimation period. The results from Jacobsen and Naug (2004:2) did also support this theory when testing their whole estimation period 1990-2004 and even when specifically testing the period during the banking crisis. The sign of the coefficient was inconsistent in the different models. We did not see any clear pattern whether the debt had a negative or positive impact on the housing prices. With no consistency, there is hard to find any economic reasoning for the different outcomes. This might be because the debt is very correlated with both the income and the housing stock (see covariance matrix in appendix). As we have discussed earlier, there is also a question whether the debt really affects the housing prices directly. We will therefore exclude all variables that include the debt.

The differenced unemployment did not comply with economic reasoning since the coefficient turned out to be positive in some models. It is not intuitive that an increasing unemployment will increase the housing prices. In other models, the coefficient was far from significant. We will therefore exclude the unemployment as a variable.

The differenced income variable became significant with a positive coefficient only when it was lagged two times. Economic theory indicates that the coefficient must be positive, since a decrease in disposable income will not cause an increase in housing prices. We cannot see any clear economic reasons for why the second lag should be different from the other lags.

In addition, including the significant second lag also affected other parameters, and turned some of these insignificant. This can be caused by high correlation with other variables. Because we want to be certain of the economic reasoning behind our model, we choose not to include the income in the short-term dynamics, as we also have argued that income might only be relevant in the long run. The cointegrating relationship between the income and the housing prices was significant in almost every model. As we have discussed earlier, we need to include all the cointegrating equations in a long-term solution to see the total impact from these variables, before we can make any conclusions regarding their economic reasoning.

It seems that the relationship between the housing stock and the income might have given some unintuitive results in the short-term dynamic. The coefficients should have opposite signs with the housing stock being negative. The problem is that this mainly applies in the long run, not in the short run. This is because the housing stock, and partly income as well, are considered as given in the short run. The insignificant and positive coefficients for housing stock could therefore be a result of the variable being estimated in a short term dynamic. The cointegration vector including income and housing stock proved to be significant and with economic intuitive signs. Since income is included in two cointegrated variables, we will check the total effect of the cointegrating equations after calculating the long-term solution.

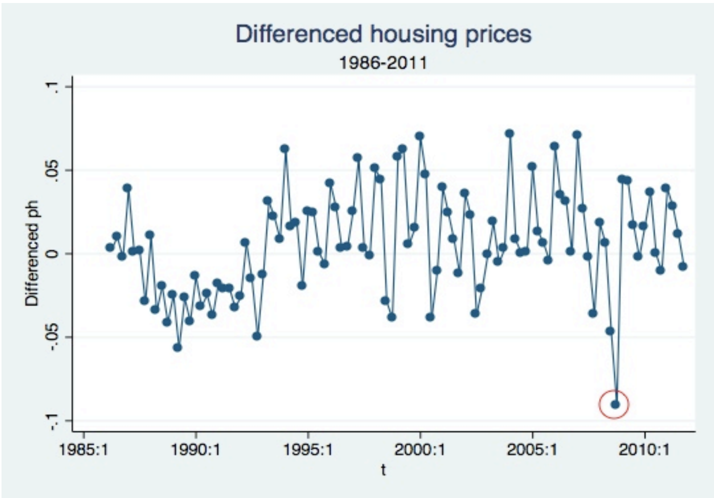
We include the first and fourth lag of the differenced housing prices. These two were significant in almost every model they were included in, and it makes sense in economic theory. The first lag implies that the change in housing prices last period has an effect on the change in housing prices this period. The fourth lag includes an effect from the change in the same quarter previous year, which makes sense if the housing prices have more seasonal variations than the other variables. This way it might catch some of the seasonal adjustment the dummy variables do not pick up.

Both real and nominal interest rates are significant in many of the models. Some lagged variants gave a positive coefficient, which would not be consistent with economic theory. The coefficient of the real interest rate lagged one period turned out to be negative and significant in most models. This complies with economic theory and therefore we choose to include the real interest rate with one lag.

The dummy variables for development in GDP did not become significant in any of our models. The coefficient is positive in most of the models, which is expected as the dummy was set to 1 in the periods with booms, but neither of the coefficients in the different models was even close to being significant. The coefficients were also very close to zero. It does not mean we can reject all hypotheses of GDP or business cycles affecting the housing prices, but the dummy variable describing whether there is a boom or recession in the economy was not found significant describing the housing prices.

The outcome regarding the seasonal dummies was inconsistent. The first quarter was always significant, the second quarter was significant in about half the models and the third one was very rarely significant. We therefore excluded the seasonal dummy for the third quarter. When we found the model we wanted to apply, we tested the first two quarters, and included both when they turned out to be significant. We also calculated the long-term solution for the different models, to check that it matches economic reasoning. This is shown for our final model later in the thesis.

Figure 30. Plot of the differenced housing prices, with 2008q4 highlighted



When we plotted the differenced data for housing prices in figure 30, we discovered an outlier in the observation from 2008q4. Often these extreme observations are due to measurement errors, but this is when the financial crisis hit Norway the hardest. A fall of approximately 9% in the housing prices is way beyond the normal. We will therefore include a dummy variable for this quarter to avoid this extreme value from making too much impact on the model.

As a result of the discussion and estimations above, we ended up with the following model:

Figure 31. Regression results of housing price model

Source	SS	df	MS
Model	.073554959	8	.00919437
Residual	.0268479	90	.00029831
Total	.100402859	98	.001024519

Number of obs = 99
F(8, 90) = 30.82
Prob > F = 0.0000
R-squared = 0.7326
Adj R-squared = 0.7088
Root MSE = .01727

D.ph	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ph						
LD.	.2133701	.076806	2.78	0.007	.0607816	.3659586
L4D.	.1838645	.0746897	2.46	0.016	.0354803	.3322486
ce_ph_yd						
L1.	-.0560789	.0104852	-5.35	0.000	-.0769097	-.0352482
ce_h_yd						
L1.	-.1714007	.0388825	-4.41	0.000	-.2486474	-.0941539
RRT						
L1.	-.7203277	.1527521	-4.72	0.000	-1.023796	-.4168589
Seas						
--.	.0368499	.0058846	6.26	0.000	.0251592	.0485406
L1.	.0164276	.0048367	3.40	0.001	.0068188	.0260365
q4_08	-.0772688	.0185108	-4.17	0.000	-.1140437	-.0404938
_cons	.0474123	.0143506	3.30	0.001	.0189022	.0759223

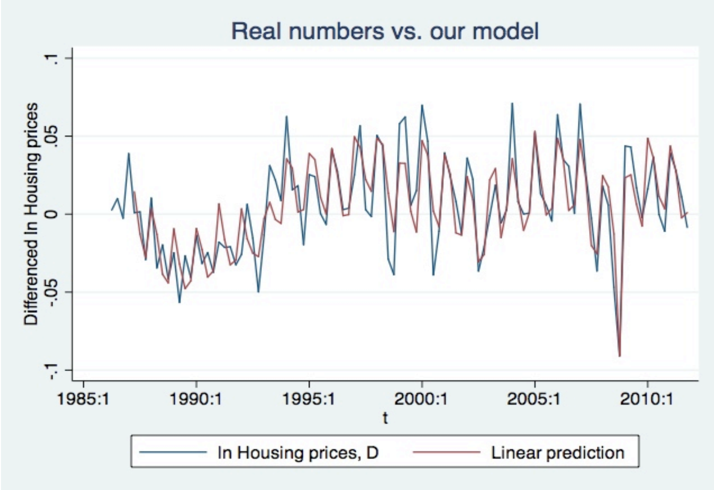
The output from figure 31, can be written as:

$$\Delta ph = \beta_1 \Delta ph_{t-1} + \beta_2 \Delta ph_{t-4} + \beta_3 ce_{ph\_yd,t-1} + \beta_4 ce_{h\_yd,t-1} + \beta_5 RRT_{t-1} + \beta_6 Seas_t + \beta_7 Seas_{t-1} + \beta_8 q4_{08} + constant$$

All the coefficients, except for Δph_{t-4}, are significant at 1% level. From the short term dynamic we can see that if the housing prices increased by 1% in the previous quarter (q_{t-1}) or the same quarter last year (q_{t-4}), our model suggests todays housing prices will increase by 0,213% or 0,184% respectively. The interest rate seems to have a great negative impact, which makes sense. The constant and seasonal dummies are positive, and imply a positive drift in the housing prices, where the impact of the drift is seasonally dependent.

If we plot our fitted value together with the actual figures, illustrated in figure 32, we see that the model explains the development in housing prices pretty well. About 70% of the fluctuations in the dependent variable can be explained by the explanatory variables. It could of course be possible to make a better fit, but we want to have economic reasoning for all our parameters.

Figure 32. The development in real housing prices vs. predication from model



In a long-term equilibrium, the variables in the short-term dynamics are equal to zero. This means all the differenced variables will be set as zero, and we will remove the constant and dummies since they do not have any impact on the equilibrium housing prices. The interest rate and the two cointegrated equations are then the variables left. The regression has estimated the coefficients, and we can use these to see how the housing prices are affected in the long run. The long-term solution can be expressed as:

$$0 = -0{,}0560789(ph + 0{,}1483129yd) - 0{,}1714007(h - 1{,}206192yd) \\ - 0{,}7203277RRT$$

The coefficient -0,0560789 in front of the cointegration vector normalized on *ph* is called the *speed of* adjustment. It describes the speed of adjustment back to equilibrium and measures the proportion of last period's equilibrium error that is corrected for (Brooks, 2008). The inverse of the speed of adjustment how many time periods it will take after a shock to get back to equilibrium. In our model that will be $\frac{1}{0{,}056} \approx 18$ quarters, or 4,5 years. The speed of adjustment can be interpreted as when the housing price is above its estimated long-term equilibrium in quarter t_{-1}, it will fall by 0,0560789% in period *t*, if everything else stays the same. The opposite if the price is below long-term equilibrium.

If we solve the equation and put ph alone on the left hand side we have the long-term solution:

$$ph = 3{,}538317065yd - 3{,}056420508h - 12{,}8448971RRT$$

We see that the model generates economically reasonable relationships for the long term, since an increasing income will increase the housing prices, while increasing housing stock or interest rates will decrease the housing prices. As we can see, the expected total income effect was not altered by the cointegrating vector ce_ph_yd, which we discussed earlier in the cointegration part. The second cointegrating equation also containing income reversed the effect. The results of our model will therefore indicate the relationships between the housing prices, disposable income, housing stock and interest rate in the long run. When the variables are on logarithmic form, the coefficient will be the approximate elasticity of the variable relative to housing prices for small changes in the variable. The changes in *percent* multiplied with the coefficient will give the percentage change in housing prices. Changes in the real interest rate after taxes will directly affect the housing prices with the coefficients value for every *percentage point* the RRT changes. The following examples illustrate the effects:

- The total disposable income increases with one percent, e.g. from 100 billion NOK to 101. That gives an increase of approximately $1\% * 3{,}54 = 3{,}54\%$ on the housing prices in the long run.
- An increase from 150BNOK to 155BNOK in the housing stock is a percentage increase of 3,33%. With all other variables kept still, that would decrease the housing prices with approximately $3{,}33\% * \ 3{,}06 = 10{,}19\%$ in the long run.
- If the real interest rate decreases by 0,5 percentage points (e.g. from 5 to 4,5% or 3 to 2,5%), the housing prices would increase with approximately $0{,}5\% * 12{,}84 = 6{,}42\%$ in the long run.

These numbers gives us an understanding of how the factors affect the housing prices in the short-run and in the long run. In the next chapter we will test whether our model is "approved" in terms of statistical validity.

10. The statistical validity of the model

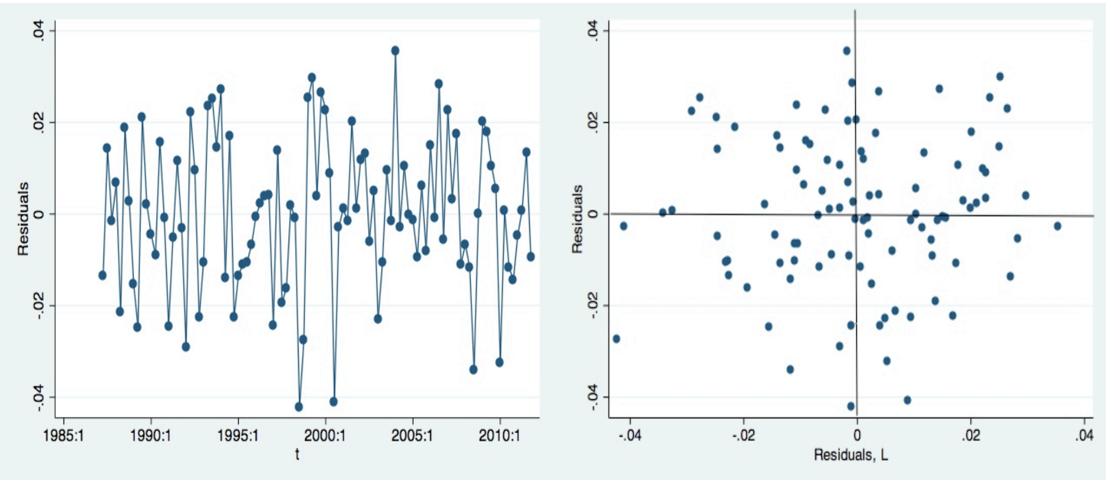
10.1 Statistical and graphical tests

Stata gives access to a large number of diagnostic tests, which we used to check the statistical analysis and the validity of our model. We use both graphical and statistical methods. All the tests we use are explained earlier in the thesis.

When estimating an error correction model, a lot of factors are sensitive to the violation of standard assumptions about the disturbances and to the exclusion of relevant independent variables. This includes the properties of the OLS estimators, parameters and standard errors, together with any associated test procedures. Therefore, we have to remember that the usefulness and validity of such tests cannot be taken for granted when the dependent variable and some of the independent variables are non-stationary (Gerrard and Godfrey, 1998). We test several aspects to make sure the standard assumptions are not being violated. All test statistics can be found in the appendix.

By looking at the residual plot in figure 33, we can see that the residuals show no clear signs of autocorrelation. There is no clear pattern in the residuals over time, the points are randomly spread and the residuals do not seem to cross the x-axis either too frequently or too little.

Figure 33. LHS Residuals over time. RHS Residuals vs. lagged residuals



The graphical tests are satisfying, but we need to support the results with statistical tests. The p-value for the Breusch-Godfrey serial correlation Lagrange multiplier test with four lags is estimated to 0, 8581, so we cannot reject the null hypothesis of no autocorrelation. This is also supported by the Ljung-box test, where we could not reject the null hypothesis of no autocorrelation of the first and fourth order.

Figure 34. LHS Histogram of residual density. RHS Standard p-plot

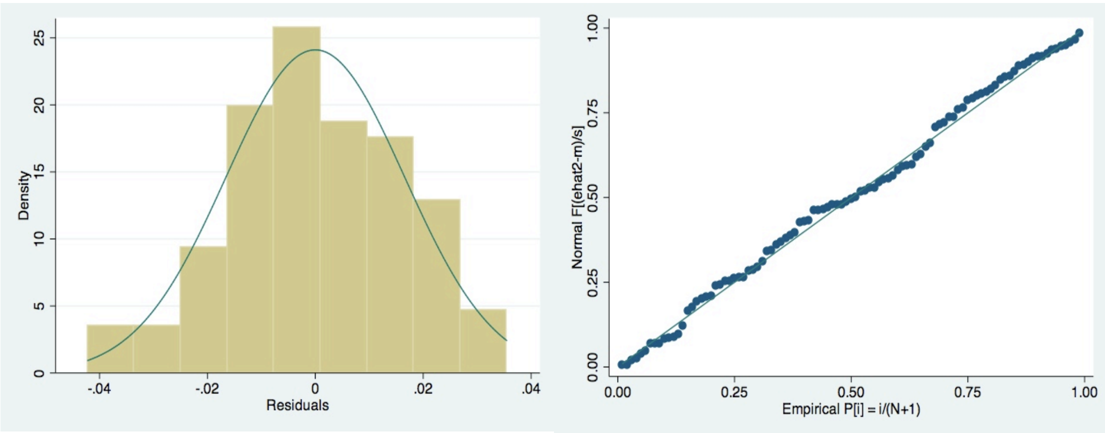


Figure 34 shows that the residuals seem to be normally distributed. The residuals are quite symmetrical, though with minor tendencies of skewness. The tail of the distribution is seemingly not fat and can be characterized as *mesokurtic*. If the hypothesized distribution describes the data sufficiently, the plotted points will approximately be a straight line. If the plotted points deviate significantly from the straight line, especially at the ends, then the hypothesized distribution is not appropriate. This is not the case, and the graphical results seem to conclude that the residuals are normally distributed. To support the evidence of normality, we test the residuals with the Jarque-Bera test for normality. The test statistic's p-value is 0,5545, and we cannot reject the null hypothesis, which says that the residuals are both symmetric and mesokurtic.

To test whether the residuals are heteroscedastic, graphical tests cannot be used. Therefore, we use the two tests presented earlier in the thesis, White's test and the ARCH test. The p-value from White's test was 0,4594. We choose to include 4 lags in the ARCH test, since we have quarterly data. The p-value from the test varies between 0,4732 and 0,6927, depending on chosen number of lags. Therefore, we cannot reject any of the null hypotheses in either of the two tests, which say that the residuals are homoscedastic.

Finally, we will test if we have chosen the best-suited functional form. An assumption in the classic linear regression model is that the appropriate function form is linear. We test this by using Ramsey's RESET test. The p-value of the test is 0,3553, and we cannot reject the null hypothesis that the functional form was correct and there are no omitted values.

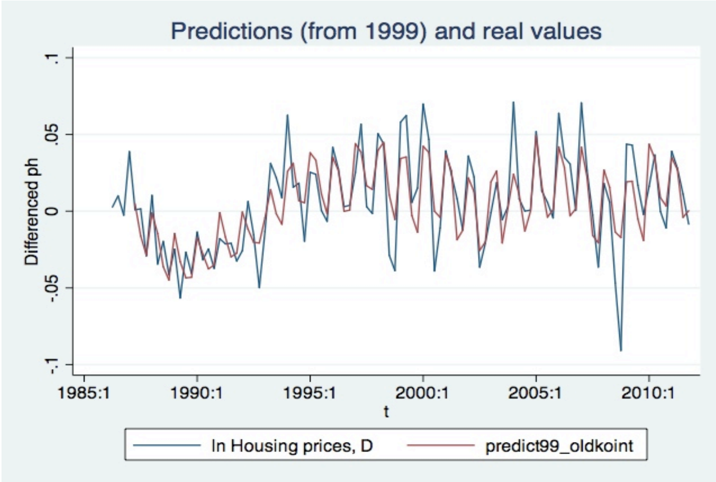
Based on the tests performed above, we believe that our model describing the development in the Norwegian housing prices, do not violate any of the standard assumptions in the statistical methods described in this thesis.

10.2 Prediction and consistency in variables

We have estimated a model that explains the housing prices. We want to be certain of our choice of variables and to check whether our economic reasoning behind the model is consistent through the estimation period. To do this, we estimated the model with the same variables only for a part of the data series. In our data series, we decided to divide the data series approximately in half. In other words, we estimated new coefficients on the basis of the data from 1986-1998(the regression output can be found in the appendix). To validate our choice of variables, we predicted the second half of the data series based on the new coefficients, and compared the prediction to the development in housing prices. If the prediction described the development well, it may be an indication of that our choice is consistent through the time series.

Figure 35 illustrates the fit between the predicted housing prices and the housing prices, when the period from 1999 to 2011 is the predicted part.

Figure 35. Predited values from 1999 to 2011 vs. housing prices



We can see that the predicted values fit the prices quite well. The adjusted R-squared is 0,63 for the entire time period. This may be a good indication that our model is robust and our variables economically grounded through the estimation period.

10.3 Criticism and deficiencies in the analysis

When working with time series and statistical analysis, we have to stay critical to any findings and results. Although our time series have all been gathered from valid and reliable sources, several of the data series have been subject of revising during the period. This can cause data in the same series to be non-comparable and the validity of the statistical analysis can be compromised.

Some of the variables are expected to be non-stationary or stationary in the analysis. These assumptions are essential for the econometric model's predictability. The tests used to determine whether the variables are stationary or not, are very sensitive to the option whether or not to include a deterministic trend and/or constant. This problem also arises when estimating the cointegrating equation. The issue is little discussed in textbooks and research papers. We have read many papers on the subject, but none has justified their choice with economic or statistical reasoning. Therefore, we have to account for any errors in our economic reasoning regarding the choice of deterministic trends and constants. To site Ahking (2001): "*In sum, we believe that the treatment of the deterministic components in a co-integration model has not been addressed adequately in the empirical literature. The* *consequences of mis-specifying the deterministic components in actual applications are largely unknown."*

Our data set contains quarterly data from 1986 to 2011, which are 104 observations. The use of lags and differenced variables made the numbers of observations even less. It is difficult to say whether there are enough observations to draw statistical conclusions, so we have to stay critical to the effects estimated by the model. We should probably have based the model on a bigger sample size to be able to draw clearer conclusions. On the other hand, the two housing models presented earlier in the thesis had both a smaller data set than us. A larger sample could also cause problems with structural changes in the markets and in the variables. Therefore, it can be difficult to decide what the optimal sample size should be. The optimal sample size would most likely depend on what kind of analysis to be conducted.

11. Conclusion

The housing prices have increased by over 300% since 1986. Except for the downturn during the banking crises and the financial crisis, there has been a continuously growth. In this thesis we have analysed the factors behind the growth in prices on the basis of an econometric model and empirical analysis.

We created an Error Correction model, which estimates the short –and long-term effects in the Norwegian housing prices. Our main findings are that the real housing prices are affected by previous housing prices, the real disposable income, the housing stock and the real interest rate after taxes.

Prices in the short term are affected by previous housing prices: the last quarter and the same quarter the previous year.

In the long-term, the housing prices are affected by the housing stock, real disposable income and the real interest rate. The housing prices will increase by approximately 3,54% for each percentage increase in disposable income, decrease by 3,06% for each percentage increase in housing stock and decrease by 12,84% for each percentage point increase in real interest rates after taxes.

We have also estimated the different phases in the business cycles in the period, but these dummy-variables did not turn out to be significant.

References

Ahking, Francis W. (2001): “*Model mis-specification and Johansen’s co-integration analysis: an application to the US money demand*”. Journal of Macroeconomics 24 (2002) 51–66. Available from http://web2.uconn.edu/ahking/Ahking02.pdf

[Downloaded 4.12.2008]

Anundsen, A. K. and Jansen, E. S. (2011): *“Self-reinforcing effects between housing prices and credit”, Evidence from Norway*, Discussion Papers No. 651, March 2011. Statistics Norway, Research Department. Available from: <www.ssb.no/publikasjoner/pdf/dp651.pdf> [Downloaded 29.08.2012]

Bakken C. (2010): “*Frisk start på boligåret”,* Nettavisen.no, 13.01.2010. Available from: <http://www.nettavisen.no/bolig/article2798766.ece> [Downloaded 15.11.2012]

Beckert, Walter (2011): *”Financial Econometrics – Lecture notes from spring 2011*”, Birkbeck University of London. Available from <http://www.ems.bbk.ac.uk/for_students/bsc_FinEcon/fin_economEMEC007U/VAR.pdf> [Downloaded 15.11.2012]

Benedictow, Andreas (2006): “*Norsk økonomi en konjunktur historie”*, Samfunnsspeilet nr. 4, 2006, Statistics Norway. Available from <http://www.ssb.no/ssp/utg/200605/14/> [Downloaded 15.09.2012]

Bjørnland, Hilde C. (2012): “*Rikdom til besvær*”, BI Business Review, 30.08.2012. Available from: <http://www.bi.no/bizreview/artikler/rikdom-til-besvar-/> [Downloaded 25.10.2012]

Bolig-abc: “*Husbankens livsløpstandard”*. Available from: <http://www.boligabc.no/bakgrunnstoff/husbankens_livslopstandard/> [Downloaded 14.11.2012]

Boug, P. and Dyvi, Y. Johansen, P. and Naug, B. (2002): “*MODAG – en makroøkonomisk modell for norsk økonomi. Sosiale og økonomiske studier 108”*, Statistics Norway. Available from: < http://www.ssb.no/emner/09/90/sos108/sos108.pdf> [Downloaded 01.09.2012]

Boug, P.and Dyvi, Y. (2008): *“MODAG - En makroøkonomisk modell for norsk økonomi”*, Sosiale og økonomiske studier 111, updated version 2008. Available from: <http://www.ssb.no/emner/09/90/sos111/> [Downloaded 01.09.2012]

Becker, Cecilie Langum (2012): “*Vi har ventet på dette lenge*”, Dagens Næringsliv. Available from <http://www.dn.no/eiendom/article2479629.ece> [Downloaded 02.10.2012]

Bjørnland Hilde C., Leif Brubakk and Anne Sofie Jore (2004): *“Produksjonsgapet i Norge – en sammenlikning av beregningsmetoder*”, Penger og Kreditt 04/2004, Norwegian Bank. Available from: < http://www.norges-bank.no/upload/import/publikasjoner/penger_og_kreditt/2000-01/produksjonsgapet.pdf > [Downloaded 28.10.2012]

Brooks, Chris (2008): “*Introductory econometrics for finance, second edition*”, Cambridge University Press, 2005.

Byggteknisk forskrift (2010): “*Forskrift 26. Mars 2010 nr. 489 om tekniske krav til byggverk”*, Lovdata. Available from: <http://lovdata.no/cgi-wift/wiftldles?doc=/app/gratis/www/docroot/ltavd1/filer/sf-20111209-1323.html&emne=rullestol*&> [Downloaded 02.10.2012]

De Rosa, I. and Horjen, H. W. (2012): “*Har ingen snarlig løsning”*, Aftenposten.no, 04.08.2012. Available from: <http://www.aftenposten.no/okonomi/innland/Har-ingen-snarlig-losning-6957437.html> [Downloaded 14.11.2012]

Dnb.no (2012): “*Priser og renter på lån for privatkunder”*, dnb.no, 15.11.2012. Available from: <https://www.dnb.no/privat/priser/laan.html> [Downloaded 15.11.2012]

DNB Eiendom Nybygg (2012): “*Vi blir flere – hva skjer med boligbehovet? “*, DNB Eiendom, 16.05.2012. Available from <http://em.dnbeiendom.no/vi-blir-flere-hva-skjer-med-boligbehovet> [Downloaded 02.10.2012]

Duke Education (2005): “*Lecture notes, Decision 411*”, Duke University, 16.05.2005. Available from: http://people.duke.edu/~rnau/411home.htm (18.11.2012)

Eika, Torbjørn (2007): *"En oljesmurt økonomi – Med harelabb over 35 års konjunkturhistorie"*, Samfunnsspeilet nr. 5-6, 2007, Statistics Norway. Available from: <http://www.ssb.no/ssp/utg/200705/15/> [Downloaded 04.09.2012]

Eika, Thorbjørn (2008): "*Det svinger i norsk økonomi*", Samfunnsspeilet nr. 5-6, 2008, Statistics Norway. Available from: < http://www.ssb.no/ssp/utg/200705/15/> [Downloaded 04.09.2012]

Eitrheim, Ø. and S. Erlandsen (2004): "*House price indices for Norway 1819-2003"*, Chapter 9 in Eitrheim, Ø., J.T. Klovland and J.F. Qvigstad (eds.), Historical Monetary Statistics for Norway 1819-2003, Norges Bank Occasional Papers no. 35, Oslo, 2004. Available from: <http://www.norges-bank.no/upload/import/publikasjoner/skriftserie/35/chapter9.pdf> [Downloaded 23.10.2012]

Elder, J and Kennedy, P. (2001): "*Testing for unit roots: What should students be taught?*", Journal of Economic education. Available from: <http://www.ufrgs.br/decon/virtuais/eco02007u/pasta/artigos/elder%20kennedy.pdf> [Downloaded 18.11.2012]

Enders, Walter (1995): "*Applied Econometric Time Series*", John Wiley & Sons, Incorporated, 1995. Available from: <http://public.econ.duke.edu/~boller/Econ.350/Enders.pdf> [Downloaded 15.11.2012]

Eurostat (2012): "*Housing statistics – Statistics Explained"* 28.08.2012. Available from: <http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Housing_statistics#Tenure_status> [Downloaded 15.11.2012]

Finansnæringens Fellesorganisasjon (2012), "*Spørsmålene som stilles*", fno.no. Available from: <http://www.fnh.no/Hoved/Aktuelt/Sporreundersokelser/Forventningsbarometeret/Forklaring-til-forventningsbarometeret/Om-forventningsbarometeret/> [Downloaded 15.11.2012]

Finn.no (2012:1): *"Bolig til salgs Bergen Sentrum"*, 14.11.2012. Available from: <http://www.finn.no/finn/realestate/homes/result?areaId=20465> [Downloaded 14.11.2012]

Finn.no (2012:2): *"Bolig til leie Bergen Sentrum"*, 14.11.2012. Available from: <http://www.finn.no/finn/realestate/lettings/result?areaId=20465> [Downloaded 14.11.2012]

Freedman, David A. (1983): "A note on screening regression equations", The American statistician, Vol 37, no 2. Available from: <http://pluto.mscc.huji.ac.il/~mszucker/REGR/freedman.pdf> [Downloaded 14.12.2012]

Gerrard, W. J. and Godfrey, L. G. (1998): "*Diagnostic checks for single-equation error-correction and autoregressive distributed lag models*", The Manchester School Vol 66 No. 2 March 1998. Available from: <http://onlinelibrary.wiley.com/store/10.1111/1467-9957.00098/asset/1467-9957.00098.pdf?v=1&t=har0gr0m&s=35d63548213a754ed4b15afe20aa784b660ec171> [Downloaded 14.11.2012]

Gjedrem, S. (2010): "*Boligfinansiering i Norge*", lecture OMF forum, 27.01.2010, Norwegian Bank. Available from: < http://www.norges-bank.no/upload/78046/figurer-27-01-2010.pdf> [Downloaded 25.09.2012]

Grytten, O. H. (2009): "*Boligboble?*", Magma, 5/2009. Available from: <http://www.magma.no/boligboble> [Downloaded 25.09.2012]

IMF, Internatinal Monetary Fund (2004): "*The Global House Price Boom*", World Economic Outlook, September 2004. Available from: <http://www.imf.org/external/pubs/ft/weo/2004/02/pdf/chapter2.pdf> [Downloaded 26.09.2012]

Jacobsen, D. H. and Naug, B. E. (2004:1): "*Hva påvirker gjeldsveksten i husholdningene?*", Penger og Kreditt, nr. 2/2004, Norwegian Bank. Available from: < http://www.norges-bank.no/upload/import/publikasjoner/penger_og_kreditt/2004-02/gjeldsveksten.pdf > [Downloaded 04.09.2012]

Jacobsen, D. H. and Naug, B. E. (2004:2): "*Hva driver boligprisene*", Penger og Kreditt, nr. 4/2004, Norwegian Bank. Available from: < http://www.norges-bank.no/upload/import/publikasjoner/penger_og_kreditt/2004-04/jacobsen.pdf > [Downloaded 01.09.2012]

Jansen, Eilev S. (2011): "*Hva driver utviklingen i boligprisene*?", Samfunnsspeilet 5-6 2011, Statistics Norway. Available from: <http://www.ssb.no/samfunnsspeilet/utg/201105/13/> [Downloaded 23.08.2012]

Johansen, Per Richard and Eika, Torbjørn (2000): "*Drivkrefter bak konjunkturforløpet på 1990-tallet*", Økonomiske analyser 6/2000, Statistics Norway. Available from: <www.ssb.no/emner/08/05/10/oa /200006/johansen.pdf> [Downloaded 17.11.2012]

Klovland, Jan Tore and Helliesen, Magnus Kvåle (2012): "*Business cycles analysis – Lecture notes from spring 2012*", Norwegian School of Economics.

Kydland, F. E. and E. C. Prescott (1990): "*Business Cycles: Real Facts and a Monetary Myth*", Federal Reserve Bank of Minneapolis, Quarterly Review. Available from: http://www.minneapolisfed.org/research/qr/qr1421.pdf [Downloaded 17.11.2012]

Sættem, Johan B (2012): "*Slik har norske boligpriser utviklet seg gjennom de siste 200 årene*", Nrk.no. Available from: http://www.nrk.no/nyheter/norge/1.8273402 [Downloaded 14.11.2012]

Ministry of Finance (2000): "*En strategi for sysselsetting og verdiskaping* ", Norges Offentlige Utredninger 2000:21. Available from: <http://www.regjeringen.no/nb/dep/fin/dok/nouer/2000/nou-2000-21/16/2/4.html?id=360364> [Downloaded 22.11.2012]

Ministry of Justice and Public Security (1994): "*Personskadeerstatning*", Norges Offentlige Utredninger 1994:20. Available from:<http://www.regjeringen.no/nb/dep/jd/dok/nouer/1994/nou-1994-20/4/1/7.html?id=374301> [17.11.2012]

Ministry of Labour (2012): "*Grunnlaget for inntektsoppgjørene 2012 4.2*", Norges Offentlige Utredninger 2012: 11. Available from: <http://www.regjeringen.no/nb/dep/ad/dok/nouer/2012/nou-2012-11/5/2.html?id=681228> [Downloaded 30.10.2012]

Romer, Christina D. (1999): “*Changes in business cycles: Evidence and explanations*”, Journal of Economic Perspectives 13, 23-44. Available from: <http://emlab.berkeley.edu/~cromer/JEP_Spring99.pdf> [15.11.2012]

Røed, K. (1996): “*Unemployment Hysteresis – Macro Evidence from 16 OECD Countries*”, Empirical Economics, Vol 21. Available from: <http://link.springer.com/article/10.1007/BF01180703?LI=true> [Downloaded 05.11.2012]

Sommervoll, Dag Einar (2007): *”Gjeldsrenter og skatt: Skattereformen av 1992 uten effekt på husholdningenes gjeld?”*, Økonomiske analyser 5/2007, Statistics Norway. Available from: <http://www.ssb.no/emner/08/05/10/oa/200705/sommervoll.pdf > [Downloaded 23.09.2012]

SSB (2012:1): “*Økonomisk utsyn*”, Økonomiske analyser 1/2012, Statistics Norway. Available from: < http://www.ssb.no/oa/tidligere_utgaver.html> [Downloaded 13.10.2012]

SSB (2012:2): “*Langvarig moderat konjunkturoppgang*”, Konjunkturtendensene for Norge og utlandet, 06.10.2012. Available from: <http://www.ssb.no/kt/> (15.11.2012)

SSB (2012:3): “*Statistikkbanken*”. Available from: <http://statbank.ssb.no/statistikkbanken/> [Downloaded 15.11.2012]

SSB (2012:4): “*Konjunkturtendensene for Norge og utlandet- Langvarig moderat konjunkturoppgang*”, Statistics Norway, 06.09.2012. Available from: <http://www.ssb.no/kt/ > [Downloaded 10.11.2012]

SSB (2010): “*Byggjekostnadsindeks, bustader – bikbol*”. Available from <http://www.ssb.no/bkibol/om.html> [Downloaded 15.11.2012]

The Financial Supervisory Authority of Norway (2011): “*Finansielt utsyn 2011*”, The Financial Supervisory Authority of Norway. Available from: <http://www.finanstilsynet.no/Global/Venstremeny/Rapport/2011/Finansielt_utsyn_2011.pdf> [Downloaded 15.11.2012]

TNS Gallup (2012): “*Om Norsk Trendindikator*”. Available from: <http://www.tns-gallup.no/?aid=9080115> [Downloaded 15.11.2012]

Wold, G. C. (2009): "*Økte byggekostnader fra 1. August*", Bonytt, 01.08.2009. Available from: <http://www.klikk.no/bonytt/boligokonomi/article486046.ece> [Downloaded 02.10.2012]

Appendix

Variable summary

Variable characteristics and correlation matrix

Variable	Mean	Std. Dev.	Min	Max
ph	-.3500847	.3302983	-.9550527	.1370843
d	14.10844	.3531193	13.66853	14.707
yd	12.23766	.166802	11.99072	12.56439
h	14.46525	.1302815	14.28913	14.69901
devigdp	-.1855444	2.552619	-4.86462	6.09063
NTI	22374.11	197429.5	-17.4056	1743670
unemps	-3.908524	.2250459	-4.3518	-3.485302
RRT	.0279251	.0160933	.000533	.07541

	ph	d	yd	h	devigdp	NTI	unemps	RRT
ph	1.0000							
d	0.9590	1.0000						
yd	0.9620	0.9852	1.0000					
h	0.9693	0.9954	0.9932	1.0000				
devigdp	0.2950	0.1803	0.1869	0.1781	1.0000			
NTI	0.1239	0.1503	0.1422	0.1418	0.1654	1.0000		
unemps	-0.6900	-0.5816	-0.6221	-0.6049	-0.6135	-0.2004	1.0000	
RRT	-0.6404	-0.6022	-0.6216	-0.6323	-0.1716	-0.1372	0.3792	1.0000

Stationarity test

We used the Augmented Dickey-Fuller Test to check the variables for stationarity. To find the appropriate number of lags, we used the Bayesian Information Criterion. This number of lags was used in the ADF test for stationarity, with the different options. The housing prices, debt and income are adjusted by the private consumption deflator. Δ in front of a variable means that it has been differenced. The results are illustrated in the following table:

Results from ADF test

Variable	# of lags	w/ constant, w/ trend	w/ constant, wo / trend	wo/ constant, wo/ trend
Ln Housing prices	5	-3,309	-	-
Ln Debt	5	-2,351	-	-
Ln Housing Stock	4	-3,530	-	-
Ln Income	4	-0,24	3,678	-
Real Interest rate	1	-4,403**	-	-
Trend indicator	0	-8,868**	-8,775**	-8,718**
Unemployment	1	-2,010	-1,212	-0,706
ΔLn Housing prices	5	-	-3,007*	-2,879**
ΔLn Debt	4	-	-4,354**	-
ΔLn Housing Stock	4	-	-2,411	-2,454*
ΔLn Income	5	-	-3,738**	-
ΔUnemployment	0	-	-11,003**	-11,044**

We have used a 5% significance level in our analysis. Values that are seen as stationary with a 5% significance level is therefore marked by one star (*), while those who are significant at 1% significance level are marked with two stars (**). Empty cells in the table means there was no point in testing with these properties according to economic theory and significance in the tests.

Cointegration

Finding the number of possible cointegrating relationships

```
. vecrank ph d yd h

                       Johansen tests for cointegration
Trend: constant                                         Number of obs =     102
Sample:  1986:3 - 2011:4                                         Lags =       2
-------------------------------------------------------------------------------
                                                       5%
maximum                                      trace  critical
  rank    parms       LL       eigenvalue  statistic  value
    0      20     1410.853          .      131.7224   47.21
    1      27     1449.9483     0.53540     53.5317   29.68
    2      32     1464.363      0.24621     24.7025   15.41
    3      35     1474.59       0.18170      4.2484    3.76
    4      36     1476.7142     0.04080
-------------------------------------------------------------------------------
```

Finding the optimal number of lags in the cointegrating relationships

```
. varsoc ph d, noconstant maxlag(5)
```

Selection-order criteria
Sample: 1987:2 - 2011:4 Number of obs = 99

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
1	499.318	.	4	.	1.5e-07	-10.0064	-9.96401	-9.90158
2	509.078	19.518	4	0.001	1.4e-07	-10.1228	-10.0379	-9.91307
3	512.656	7.156	4	0.128	1.4e-07	-10.1143	-9.98698	-9.79969
4	520.054	14.797	4	0.005	1.3e-07	-10.1829	-10.0132	-9.76349
5	555.721	71.334*	4	0.000	6.8e-08*	-10.8226*	-10.6105*	-10.2984*

Endogenous: ph d
Exogenous:

```
. varsoc ph yd, noconstant maxlag(5)
```

Selection-order criteria
Sample: 1987:2 - 2011:4 Number of obs = 99

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
1	423.745	.	4	.	7.1e-07	-8.47969	-8.43727	-8.37484
2	460.671	73.852	4	0.000	3.7e-07	-9.14487	-9.06002	-8.93516
3	479.104	36.866	4	0.000	2.7e-07	-9.43644	-9.30917	-9.12188
4	510.585	62.962	4	0.000	1.6e-07	-9.99161	-9.82192	-9.5722
5	522.68	24.191*	4	0.000	1.3e-07*	-10.1552*	-9.94303*	-9.63089*

Endogenous: ph yd
Exogenous:

```
. varsoc h yd, noconstant maxlag(5)
```

Selection-order criteria
Sample: 1987:2 - 2011:4 Number of obs = 99

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
1	703.514	.	4	.	2.5e-09	-14.1316	-14.0892	-14.0267
2	840.191	273.36	4	0.000	1.7e-10	-16.8119	-16.7271	-16.6022
3	830.429	-19.524	4	.	2.2e-10	-16.5743	-16.4683	-16.3122
4	869.958	79.058	4	0.000	1.1e-10	-17.2921	-17.1436	-16.9251
5	876.216	12.516*	4	0.014	9.7e-11*	-17.3781*	-17.2084*	-16.9587*

Endogenous: h yd
Exogenous:

Cointegrating vector between housing prices and debt

og likelihood =	555.6443			HQIC		= -10.63979
et(Sigma_ml) =	4.57e-08			SBIC		= -10.34325

quation	Parms	RMSE	R-sq	chi2	P>chi2
_ph	9	.02369	0.5105	93.86567	0.0000
_d	9	.011096	0.6639	177.7641	0.0000

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
D_ph						
_ce1						
L1.	.0086448	.0042656	2.03	0.043	.0002843	.0170052
ph						
LD.	.2477536	.0902847	2.74	0.006	.0707988	.4247083
L2D.	-.092419	.0940268	-0.98	0.326	-.2767082	.0918702
L3D.	.0834951	.0897726	0.93	0.352	-.092456	.2594463
L4D.	.6254029	.0897623	6.97	0.000	.449472	.8013338
d						
LD.	.1349944	.2099833	0.64	0.520	-.2765654	.5465542
L2D.	-.4387741	.2092789	-2.10	0.036	-.8489531	-.0285951
L3D.	.1298262	.2139018	0.61	0.544	-.2894135	.549066
L4D.	-.493433	.2141345	-2.30	0.021	-.913129	-.073737
D_d						
_ce1						
L1.	.0035341	.0019979	1.77	0.077	-.0003816	.0074498
ph						
LD.	.0721004	.0422861	1.71	0.088	-.010779	.1549797
L2D.	.0126623	.0440388	0.29	0.774	-.0736523	.0989768
L3D.	.0756093	.0420463	1.80	0.072	-.0068	.1580186
L4D.	.0133412	.0420415	0.32	0.751	-.0690586	.095741
d						
LD.	.1669867	.0983488	1.70	0.090	-.0257733	.3597467
L2D.	-.0340201	.0980188	-0.35	0.729	-.2261334	.1580932
L3D.	.0533322	.100184	0.53	0.594	-.1430248	.2496893
L4D.	.4109827	.100293	4.10	0.000	.214412	.6075535

Cointegrating equations

Equation	Parms	chi2	P>chi2
_ce1	1	14.40338	0.0001

Identification: beta is exactly identified

Johansen normalization restriction imposed

beta	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_ce1						
ph	1	.	.	.	.	.
d	.0932852	.0245799	3.80	0.000	.0451094	.1414609

Cointegrating vector between housing prices and income

ector error-correction model

ample:	1987:2 - 2011:4	No. of obs	= 99
		AIC	= -10.16867
og likelihood =	522.349	HQIC	= -9.967154
et(Sigma_ml) =	8.96e-08	SBIC	= -9.670614

quation	Parms	RMSE	R-sq	chi2	P>chi2
_ph	9	.024238	0.4876	85.64683	0.0000
_yd	9	.013605	0.7628	289.4005	0.0000

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_ph						
_ce1						
L1.	-.0017273	.003781	-0.46	0.648	-.0091379	.0056832
ph						
LD.	.3163231	.0934891	3.38	0.001	.1330878	.4995584
L2D.	-.1112687	.0986029	-1.13	0.259	-.3045268	.0819894
L3D.	.0829353	.1006205	0.82	0.410	-.1142773	.280148
L4D.	.4583485	.0945113	4.85	0.000	.2731097	.6435874
yd						
LD.	.0493045	.1878246	0.26	0.793	-.318825	.417434
L2D.	.358187	.2205187	1.62	0.104	-.0740217	.7903956
L3D.	.0966086	.2204691	0.44	0.661	-.3355028	.52872
L4D.	.0229918	.1888502	0.12	0.903	-.3471479	.3931314
D_yd						
_ce1						
L1.	.0149734	.0021223	7.06	0.000	.0108139	.019133
ph						
LD.	-.0435368	.0524757	-0.83	0.407	-.1463873	.0593137
L2D.	.11433	.0553461	2.07	0.039	.0058536	.2228063
L3D.	-.0125694	.0564786	-0.22	0.824	-.1232655	.0981266
L4D.	-.0489092	.0530495	-0.92	0.357	-.1528842	.0550659
yd						
LD.	-.8792006	.1054265	-8.34	0.000	-1.085833	-.6725685
L2D.	-.6573067	.1237778	-5.31	0.000	-.8999067	-.4147067
L3D.	-.6216198	.1237499	-5.02	0.000	-.8641652	-.3790745
L4D.	.0192955	.1060022	0.18	0.856	-.1884649	.2270559

Cointegrating equations

Equation	Parms	chi2	P>chi2
_ce1	1	95.17115	0.0000

Identification: beta is exactly identified

Johansen normalization restriction imposed

beta	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_ce1						
ph	1	.	.	.	.	.
yd	.1483129	.0152029	9.76	0.000	.1185158	.17811

Cointegrating vector between housing capital and income

```
. vec h yd, trend(none) lags(5)

Vector error-correction model

Sample:  1987:2 - 2011:4                        No. of obs      =        99
                                                AIC             = -17.71012
Log likelihood =   895.651                      HQIC            = -17.50861
Det(Sigma_ml)  =  4.75e-11                      SBIC            = -17.21207
```

Equation	Parms	RMSE	R-sq	chi2	P>chi2
D_h	9	.00055	0.9913	10256.9	0.0000
D_yd	9	.013787	0.7564	279.4528	0.0000

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
D_h						
_ce1 L1.	-.0028035	.0006475	-4.33	0.000	-.0040726	-.0015345
h LD.	.9472984	.0939707	10.08	0.000	.7631193	1.131478
L2D.	-.2767377	.1244478	-2.22	0.026	-.5206509	-.0328245
L3D.	.2733	.1239184	2.21	0.027	.0304245	.5161755
L4D.	-.0125489	.0873642	-0.14	0.886	-.1837797	.1586818
yd LD.	-.0086533	.004585	-1.89	0.059	-.0176398	.0003331
L2D.	-.0144399	.005119	-2.82	0.005	-.0244729	-.0044069
L3D.	-.0184173	.0049784	-3.70	0.000	-.0281748	-.0086597
L4D.	-.0172736	.004293	-4.02	0.000	-.0256877	-.0088596
D_yd						
_ce1 L1.	-.0654821	.0162168	-4.04	0.000	-.0972664	-.0336978
h LD.	-1.995423	2.353606	-0.85	0.397	-6.608407	2.617561
L2D.	.8692563	3.116941	0.28	0.780	-5.239836	6.978348
L3D.	4.583986	3.103681	1.48	0.140	-1.499117	10.66709
L4D.	-2.501356	2.188139	-1.14	0.253	-6.790031	1.787318
yd LD.	-.94546	.1148363	-8.23	0.000	-1.170535	-.720385
L2D.	-.7377214	.128211	-5.75	0.000	-.9890104	-.4864324
L3D.	-.7105214	.1246906	-5.70	0.000	-.9549106	-.4661323
L4D.	.0006579	.107522	0.01	0.995	-.2100814	.2113973

Cointegrating equations

Equation	Parms	chi2	P>chi2
_ce1	1	114398.2	0.0000

Identification: beta is exactly identified

Johansen normalization restriction imposed

beta	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_ce1						
h	1	.	.	.	.	.
yd	-1.206192	.0035662	-338.23	0.000	-1.213182	-1.199202

Tests for Heterescedasticity

White's test

```
. imtest, white

White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedasticity

         chi2(33)     =    32.83
         Prob > chi2  =   0.4755

Cameron & Trivedi's decomposition of IM-test
```

Source	chi2	df	p
Heteroskedasticity	32.83	33	0.4755
Skewness	8.60	8	0.3771
Kurtosis	0.84	1	0.3605
Total	42.27	42	0.4594

ARCH-test

```
. estat archlm, lags(1,2,3,4)
LM test for autoregressive conditional heteroskedasticity (ARCH)
```

lags(p)	chi2	df	Prob > chi2
1	0.514	1	0.4732
2	0.880	2	0.6439
3	2.197	3	0.5325
4	2.235	4	0.6927

H0: no ARCH effects vs. H1: ARCH(p) disturbance

Test for autocorrelation

Breusch-Godfrey test

```
. estat bgodfrey, lags(4)

Breusch-Godfrey LM test for autocorrelation
```

lags(p)	chi2	df	Prob > chi2
4	1.319	4	0.8581

H0: no serial correlation

Ljung-Box test

```
. wntestq ehat2, lags(4)

Portmanteau test for white noise

 Portmanteau (Q) statistic =     1.0723
 Prob > chi2(4)            =     0.8986
```

Test for normality

Jarque-Bera test

```
. sktest ehat2

                    Skewness/Kurtosis tests for Normality
                                                          ------ joint ------
    Variable |        Obs  Pr(Skewness)  Pr(Kurtosis)  adj chi2(2)  Prob>chi2
-------------+---------------------------------------------------------------
       ehat2 |         99     0.3244        0.6688         1.18        0.5545
```

Test for functional form

Ramsey's RESET test

```
. estat ovtest

Ramsey RESET test using powers of the fitted values of D.ph
       Ho:  model has no omitted variables
                 F(3, 87) =      1.10
                 Prob > F =      0.3553
```

Prediction and consistency in variables

Regression results of housing price model with data from 1986-1998

```
. regress D.ph LD.ph L4D.ph L.ce_ph_yd L.ce_h_yd L.RRT Seas L.Seas

      Source |       SS       df       MS              Number of obs =      51
-------------+------------------------------           F(  7,    43) =   13.92
       Model |  .033143327     7  .004734761           Prob > F      =  0.0000
    Residual |   .01462578    43  .000340134           R-squared     =  0.6938
-------------+------------------------------           Adj R-squared =  0.6440
       Total |  .047769107    50  .000955382           Root MSE      =  .01844

------------------------------------------------------------------------------
        D.ph |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ph |
         LD. |   .1621254     .13501     1.20   0.236    -.1101483    .4343991
        L4D. |   .2945133    .128956     2.28   0.027     .0344487    .5545779
             |
    ce_ph_yd |
         L1. |   -.057842   .0207203    -2.79   0.008    -.0996285   -.0160555
             |
     ce_h_yd |
         L1. |  -.1918442   .0677252    -2.83   0.007     -.328425   -.0552633
             |
         RRT |
         L1. |  -.5154045   .2753881    -1.87   0.068    -1.070777    .0399685
             |
        Seas |
         --. |   .0263574   .0084201     3.13   0.003     .0093767    .0433381
         L1. |   .0138306   .0068676     2.01   0.050    -.0000193    .0276805
             |
       _cons |   .0390145    .034905     1.12   0.270    -.0313782    .1094072
------------------------------------------------------------------------------
```