Housing Construction in Stockholm

Fundamental Factors’ Impact on Construction Volume

Lovisa Onsbring
Abstract

Housing availability plays a central role for the mobility in the labour market and for the economic growth in a country. Housing construction in Sweden is now increasing strongly after a period of low construction during and after the financial crisis in 2008. Despite this, many of the Swedish municipalities indicate that there is a housing shortage, especially in the metropolitan regions. The main reason for this is that the population has grown at a faster pace compared to the number of homes.

This thesis uses a regression analysis to analyse what fundamental economic factors that affected housing construction in Stockholm County between 1992 and 2016.

The result from the regression analysis shows that the repo rate, inflation, GDP, population, unemployment, housing prices, and construction cost, with a lag of seven, seven, zero, one, six, two, and three quarters respectively, have a significant effect on housing construction. Furthermore, the regression model showed a low significance level for income, thus this variable was excluded from the model. Moreover, a time trend was included in the regression model. The time trend shows statistically significance, and its β coefficient implies an approximate 3% increase in housing construction per year, on average.

Moreover, the regression model, presented in this thesis, has an R-squared value of 0.73, which indicates that the model can explain 73% of the variation in housing construction in Stockholm County between 1992 and 2016.
Sammanfattning


Resultatet av regressionsanalysen visar att reporäntan, inflationen, BNP, befolkningsmängd, arbetslöshet, bostadspriser och byggkostnader, med sju, sju, noll, en, sex, två, respektive tre kvartals förskjutning, hade en signifikant inverkan på bostadsbyggandet. Vidare visade regressionsmodellen en låg signifikansnivå för variabeln inkomst, vilket medförde att denna variabel uteslöts från modellen. Dessutom inkluderades en tidsvariabel i regressionsmodellen. Tidsvariabeln visar statistisk signifikans och dess koefficient indikerar att bostadsbyggandet i genomsnitt ökade med ca 3% per år.

Regressionsmodellen som presenteras i denna uppsats har ett $R^2$-värde på 0,73 vilket indikerar att modellen förklarar 73% av variationen i bostadsbyggande i Stockholms län mellan 1992 och 2016.
Acknowledgements

First, I would like to thank my supervisor Sviatlana Engerstam for her valuable guidance, enthusiasm and constructive feedback throughout the whole process of writing this thesis. Furthermore, I would also like to express my great gratitude to Han-Suck Song, at the department of Real Estate and Construction Management at KTH, for all the valuable discussions, and for sharing his knowledge, as well as giving his input on the econometrics of this study.
# Table of contents

1. Introduction .................................................................................................................. 1
   1.1 Background ............................................................................................................. 1
   1.2 Purpose .................................................................................................................... 3
   1.3 Research question .................................................................................................. 4
   1.4 Delimitations .......................................................................................................... 4
2. Overview of the Swedish housing market ...................................................................... 5
   2.1 The role of policy regulations ................................................................................. 5
       2.1.1 Government regulations ................................................................................ 5
       2.1.2 The planning process .................................................................................... 5
   2.2 Stockholm housing market ..................................................................................... 7
3. Literature review ........................................................................................................... 9
   3.1 Business and real estate cycles .............................................................................. 9
   3.2 Behaviour based theories ....................................................................................... 13
   3.3 The q theory of housing investment ..................................................................... 13
   3.4 Housing bubbles .................................................................................................... 14
   3.5 The DiPasquale-Wheaton model .......................................................................... 15
4. Methodology and research design ............................................................................... 17
   4.1 Literature review as a theoretical basis for the research ..................................... 17
   4.2 Quantitative method .............................................................................................. 18
   4.3 Regression analysis ............................................................................................... 18
       4.3.1 Linear regression models .............................................................................. 18
       4.3.2 Distributed lag models ............................................................................... 19
       4.3.3 Time trend ..................................................................................................... 19
   4.4 Statistical tests of the data and models ................................................................ 20
       4.4.1 Stationarity ..................................................................................................... 20
       4.4.2 Correlation ..................................................................................................... 20
       4.4.3 Multicollinearity ............................................................................................ 20
       4.4.4 Autocorrelation ............................................................................................. 21
       4.4.5 Heteroscedasticity ......................................................................................... 21
5. Data collection and hypotheses .................................................................................... 22
   5.1 Variables .................................................................................................................. 23
       5.1.1 The dependent variable ................................................................................. 23
5.1.2 Independent variables and hypotheses ........................................... 23
6 Results .................................................................................................... 28
  6.1 Regression outcomes .......................................................................... 28
  6.2 Hypothesis testing ............................................................................ 31
7 Analysis .................................................................................................. 32
  7.1 Analysis of variables ......................................................................... 32
  7.2 Limitations ......................................................................................... 35
8 Conclusion ............................................................................................... 37
  8.1 Concluding remarks .......................................................................... 37
  8.2 Suggestions for further research ......................................................... 37
9 References ............................................................................................... 38
Appendices ................................................................................................. 41
  Appendix I. Regression model ................................................................. 41
  Appendix II. Statistical testing ................................................................. 42
    Test for Stationarity ............................................................................. 42
    Test for Correlation ............................................................................ 44
1 Introduction

This chapter serves as an introduction to the research subject. The chapter presents the background of the thesis together with the purpose, research question and delimitations of the thesis.

1.1 Background

There are about 4.8 million apartments in Sweden. In a short-term perspective, the number of apartments available in Sweden is constant, and due to the immobility of buildings, these apartments are not always in the best geographical place for optimal utilization. Whether there is a housing shortage or not in Sweden, depends on how we utilize the existing stock, as well as the rate that we build new and demolish existing apartments. (Boverket, 2011)

Due to increasing income, low interest rate, low amortization rate and a high population growth, the demand for housing has increased significantly in recent years. (Boverket, 2014a) Over the last two years, there has also been a significant increase in housing construction. According to the Swedish National Board of Housing, Building and Planning (Boverket), new constructions of approximately 67,000 homes begun in 2016. This is an increase of more than 30 percent compared to 2015. (Boverket, 2017b)

Housing availability is important in many aspects. It is necessary for the labour market to work and for the economic growth of society to continue. The future need for housing, and where this need is located, is therefore of great importance to the economy. Housing investments based on short-term market conditions can be inaccurate in a longer perspective due to the long life and the geographical immobility of buildings. It is therefore important to make long-term estimations of housing needs. (Boverket, 2015)

Moreover, the rapid urbanization in Sweden after World War II resulted in an urgent need for housing in the cities. This lead to the so-called Million Programme, initiated by the Swedish government. The aim of the program was to build one million homes during a decade, to overcome the current housing shortage, and to increase the standard of living. In the years 1965-1974, about 1,006,000 homes were built in Sweden. After the program was completed, housing construction
slowed down during a ten-year period, after which it began to rise again in the mid-1980s. (Boverket, 2014b) Figure 1 illustrates the housing development in Sweden between 1950 and 2016.

Furthermore, the development in the housing and real estate market has through history played a leading role in economic crises, where the Swedish crisis in the early 1990s, and the subprime mortgage crisis in the U.S. between 2007-2010 are two examples. Heavy drops in real estate prices have in many countries been associated with major disruption in the country’s economy. Usually these price drops have been preceded by a long period of rising house prices and, in many cases, an increase in the indebtedness in the household sector.

However, the housing prices in Sweden did not fall to the same extent as they did in many other countries during the financial crisis in 2008. One potential reason to this is that housing supply has been low in relation to demand, partly due to the low housing construction in Sweden since the early 1990s.

In the early 1990s, major changes to tax policy were made, which had a negative effect on housing construction. The combined effect of changes in the VAT system, raised property tax, reduced subsidies, sharply increased real interest rates after tax and a deep recession significantly reduced the demand for new housing. (Sveriges Byggindustrier, 2015)

![Figure 1. Housing construction in Sweden 1950-2016. (Sveriges Byggindustrier, 2016)](image-url)
Moreover, in a housing market survey from 2017, conducted by the Swedish National Board of Housing, Building and Planning, many municipalities indicated an imbalance in the local housing market. 255 out of a total of 290 municipalities in Sweden estimated that there is a deficit of housing in their municipality. This is an increase in housing deficit compared to the year before. According to the survey, only one municipality estimates an excess in housing. The remaining 34 municipalities considered the housing market to be in balance.

In the survey, the Swedish municipalities also rated what they consider to be the three main barriers to increase housing construction. Results from the survey showed that high production cost was seen as the main barrier. Also, lack of detail plan on attractive land and difficulties for individuals to get loans, was pointed out as limiting factors for housing construction. (Boverket, 2017a)

1.2 Purpose

The overall purpose of this thesis is to analyse what fundamental economic factors that affect housing construction in Stockholm. There are many different factors that might affect housing construction. A more market driven demand for housing in Sweden can be seen after 1992, which indicates that housing construction might be explained by fundamentals, to a larger extent, after 1992. However, the planning system and some other policy regulations in Sweden still influence the development on the housing market. Municipalities have a strong control over housing construction, and it is therefore of great importance to evaluate if municipalities are in line with economic development and housing demand. Housing shortage is a huge problem in some regions in Sweden, especially in Stockholm. To reduce problems with housing shortage, the underlying causes for housing construction must be investigated.

This thesis provides deeper insight into how fundamental economic factors affect the level of housing construction in Stockholm. This is the first study of this kind and it contributes to the research gap that is found in the literature review. In this thesis, fundamental economic factors are factors that are in line with economic theory and that are valid in several previously performed empirical studies.

The results presented in the thesis might be used by municipalities to achieve a better understanding of the impact of economic variables on housing supply, which could help them to improve their planning policies.
1.3 Research question

The thesis aims to answer the following research question:

- Can housing construction in Stockholm be explained by fundamental economic factors, and if, to what extent?

1.4 Delimitations

The thesis is delimited to analyse fundamental factors that are quantifiable, meaning that qualitative factors are excluded from the analysis in this thesis. However, political factors, such as governmental reforms, and the structure of the Swedish planning system, are briefly described in the following chapter, together with relevant literature on the subject. However, the political factors are not included in the regression analysis, although they may have a major impact on housing construction in Stockholm.
2 Overview of the Swedish housing market

This chapter presents some of the qualitative factors that may affect housing construction in Sweden. Political influences, such as governmental reforms, and the structure of the Swedish planning system, are described together with relevant literature. Furthermore, the chapter also provides a description of the housing market in Stockholm; its development over the last decades, and the current market situation.

2.1 The role of policy regulations

2.1.1 Government regulations

Lind (2003) argues that the main reason for the large variation in housing construction, seen in Sweden over time, is the Swedish housing policy. This involves, for example, changes in taxes, and the introduction of government subsidies for housing construction. (Lind, 2003)

The large decline in housing construction in the beginning of the 1990s was partly due to the financial crisis in Sweden and the recession that followed. However, also the major changes in economic policy that resulted in a more market driven demand for housing investment played an important role in the decline in housing construction. Changes in connection to the tax reform in 1993 affected the housing market. A decline in interest subsidies and interest rate guarantees for construction projects meant that the government-subsidised secondary mortgages disappeared, leading to a market with similar funding conditions for both public and private housing companies. (Emanuelsson, 2015)

2.1.2 The planning process

In Sweden, municipalities have a monopoly on spatial planning, meaning that the municipality decides how the land will be used and to what extent it will be built within the municipality. Because of the municipal planning monopoly, the municipality serves as a key actor for housing development.

Furthermore, the planning process in Sweden is regulated in the Planning and Building Act (PBL). All municipalities must have a comprehensive plan that provides guidance on how the land and water areas, and the existing urban environment should be used, preserved and development in the
long term. The comprehensive plan should cover the entire municipality. However, this plan is not legally binding. (Boverket, 2016)

The detailed development plan, on the other hand, is legally binding, and it serves as the implementation instrument for the municipality. This plan can only be produced and adopted by the municipality itself, and it regulates how land and water within an area in the municipality should be used and built. For example, requirements for the execution of buildings can be set in the plan, with details such as how large or tall a building may be, the materials to be used, or the distance between buildings. This allows the municipality to control all housing development taking place within the municipality.

The implementation period for a detailed development plan must be set to a minimum of five years and a maximum of 15 years, meaning that the rights that the plan accord are economically protected during this period. (Boverket, 2014c)

The planning process in Sweden has many times been pointed out as complicated and time consuming. There are many requirements that need to be fulfilled in order to adopt a new detailed development plan. However, a change in law in 2015 aimed to simplify the planning process in Sweden. What was previously called “simple planning procedure” is now the norm and an extended planning procedure applies under certain circumstances, for example, when the detailed development plan is not compatible with the comprehensive plan, or when the plan can be assumed to cause a significant environmental impact.

For the municipality to adopt a detailed development plan through a standard procedure, the municipality must first draw up a planning proposal, then consulted about the proposal, and finally had it reviewed. The time required for each step in this process varies, which make the planning process unpredictably. (SFS 2010:900)

Furthermore, Mayer and Somerville (2000a) examined the relationship between land use regulation and housing construction. By using quarterly data from 44 U.S. metropolitan areas between 1985 and 1996, they found that cities with more extensive regulations have a lower level of new construction. The mean construction start was found to be up to 45 percent lower in these cities. Further, their study also shows that metropolitan areas with greater land use regulations experience lower price elasticities compared to cities with less regulations. However, Mayer and Somerville divided regulation into three different categories, depending on whether they add explicit costs, uncertainty, or delays to the development process. Their results, in accordance with several other studies, showed that regulations that lengthen or causing uncertainty to the development process
have larger and more significant effects on new construction than development fees and other regulations that add costs to the development process. (Mayer and Somerville, 2000a)

In addition, several studies have found a correlation between the complexity of the planning process, the level of new construction, and the prices of new constructions. Luger and Temkin (2000) compared states in the U.S., and found that strict regulation, low level of housing construction, and high housing prices correlated. Furthermore, Glaeser and Gyourko (2002) and Boulhouwer and deVries (2002) found a similar relationship for the U.S. and Holland, respectively.

2.2 Stockholm housing market

Stockholm County has a population of almost 2.3 million people, spread over approximately one million homes. About 40 percent of the Swedish housing stock consists of tenant owned apartments (bostadsrätter), 36 percent of rental apartments (hyresrätter), and 24 percent are ownership apartments (äganderätter). (Länsstyrelsen Stockholm, 2017)

Stockholm is one of the fastest growing regions in Europe, with a population growth of over 320,000 people during the last ten years. The population growth is expected to continue, and in 2030, the county is estimated to have more than 2.6 million inhabitants. (Länsstyrelsen Stockholm, 2017)

To cope with the increasing need for housing, the region needs a high housing construction combined with increased mobility on the housing market. There is also a need for greater variety of homes being built in order to meet many different needs on the market.

The housing construction in Stockholm fell drastically in the beginning of the 1990s as a result of the financial crisis and the deregulation on the housing market. The low construction remained until 2013. However, in recent years, housing construction has increased sharply. From 2013 to 2016, Stockholm had almost 62,000 housing construction starts, including reconstruction, in the county, which can be compared to 29,000 starts in the previous four-year period. In 2016, the regional yearly target of 16,000 new homes was exceeded for the first time, when 16,250 new homes were built. However, despite increased housing construction, housing shortage increases as population in the county grows at a faster pace than the housing construction. (Länsstyrelsen Stockholm, 2017)

Moreover, in recent years, many developers in Stockholm have focused on a niche with more expensive housing. This has created an imbalance in the housing stock, resulting in difficulties for households with lower income to enter the housing market. Some consider the more expensive housing segment to be more or less saturated, while there is still a huge demand for cheaper housing and rental apartments in Stockholm. (Dagens industri, 2017)
Moreover, the housing shortage in Stockholm can have a negative effect on the labour market. Without housing for employees, the opportunities for companies to grow are limited. This may lead to companies moving to other cities.

Furthermore, in order to meet the needs of a growing population in Stockholm, the public transport system needs to be expanded. The subway system in Stockholm is under development. 19 km of extended tracks, together with eleven new stations, will provide increased accessibility within the Stockholm region. The development will create conditions for increased housing development, and by agreements, the municipalities have undertaken to construct approximately 78,000 homes until 2030, within the metro area. Moreover, an expended and capacity-efficient railroad will increase the opportunities for commuting in the county, which could facilitate housing development in new areas.

Moreover, according to Newsec (2017), the demand for new housing has grown sharply in recent years, partly due to increased immigration and urbanization. However, signs of satiation are now shown. As shortage of labour and land leads to rising construction costs, and as interest rates begin to rise, housing construction is expected to level out during 2018. Moreover, the longstanding trend of rising housing prices in Stockholm has recently turned, meanwhile a number of regulations and tax rises concerning the property sector are being discussed. These rules and regulations are causing concerns, and they are likely to impact the housing and real estate market. Rules about households’ indebtedness leads to sharper lending regulations, which will increase the cost of mortgage. (Newsec, 2017)
3 Literature review

This chapter presents a summary of existing literature on real estate cycles followed by housing supply and demand determinants. Moreover, the chapter includes the current knowledge together with theoretical and methodological contributions to the subject area.

3.1 Business and real estate cycles

Burns and Mitchell (1946) examined cyclical fluctuations of the market economic system. In their book “Measuring business cycles” they provide the following definition of business cycles:

“Business cycles are a type of fluctuation found in the aggregate economic activity of nations that organize their work mainly in business enterprises: a cycle consists of expansions occurring at about the same time in many economic activities, followed by similarly general recessions, contractions, and revivals which merge into the expansion phase of the next cycle; in duration, business cycles vary from more than one year to ten or twelve years; they are not divisible into shorter cycles of similar characteristics with amplitudes approximating their own.” (Burns and Mitchell, 1946, p. 3)

One of the main reasons for studying business cycles is to identify and understand the unstable nature of the market mechanisms. This also applies to real estate cycle analysis. (Reed et. al., 2010)

Wheaton (1999) defines a real estate cycle as follows:

“... a real estate cycle involves repeated oscillations of a market, as it continually overshoots and then undershoots its own steady state ... real estate cycles are defined as some degree of instability in the market whereby a single economic shock leads the market to oscillate around its steady state for some number of iterations” (Wheaton, 1999)

Fanning (2005) investigated the fundamental causes of business and real estate cycles. He also studied the interaction of the national business cycle and local real estate cycles, and how the market adjusts to changes in supply and demand.
Fanning (2005) categorises the real estate cycle into four phases:

**Expansion** – A period with increasing occupancy and absorption rates, rising prices and eventually an upturn in construction activity.

**Slowdown, peak and downturn** – During this period, absorption rates start to decline, but are still positive, and construction activity begins to fall. The already high real estate prices continue to rise, while occupancy remains at a high level. When the peak of the cycle is reached, absorption rates and construction activities start to decline.

**Contraction** – A period when occupancy and absorption rates decline, and prices for real estate fall. This situation eventually leads to a decline in construction activity.

**Slowed contraction, trough, and upturn** – During this period, the downward trend of construction activity reverses, and absorption and occupancy rates stabilize. After reaching the lowest point of the cycle, the real estate market commences a moderate upturn, leading to the next expansion phase.

Furthermore, two different types of real estate cycles are often distinguished; the physical cycle and the financial cycle. The physical cycle describes fluctuations in construction activity, which is affected by supply and demand, while the financial cycle, describes fluctuations in real estate prices, which is affected by capital flows. Both cycles are related to each other, as well as to other fluctuations, e.g. in vacancy levels and rent levels. (Rics, 2015)

Long term real estate cycles are primarily affected by the national changes in population and income. A growing population and a rising income lead to an increase in demand for real estate. In the short term, however, the availability of credit or capital, and the level of interest rates are the main determinants of real estate demand. Changes in interest rates affect the price paid for capital. Since real estate usually requires a large amount of debt financing, the demand for real estate is highly affected by changes in interest rates. When interest rates decline, the demand for real estate increases. (Fanning, 2005)

Previous research indicates that construction lags in relation to demand changes for space is the primarily reason for cyclical behaviour on the real estate market, identified in the long run. Changes in demand is mainly driven by fluctuations in business activity. Barras (1994) shows how each step in the real estate cycle can be linked to the parallel business cycle, and how the fundamental structure of the real estate market, with building booms and recessions, are affected by real and money economy in the business cycle. This relationship is illustrated in figure 2.
Moreover, time lags in the real estate cycle occur as the market attempts to adjust to changes in supply and demand. When there is an excess in demand, prices and rents begin to rise before the construction of new real estate starts to accelerate. Excess in supply also affect the real estate market. After a period with rising prices and rents, construction starts to accelerate, leading to excess space. Even when the excess in supply is recognized, buildings that are currently under construction still have to be completed. The construction lag, resulting from the time it takes to complete the building process, leads to an even larger excess in supply at the completion of the new buildings, since more stock has been added to the market. (Fanning, 2005) Figure 3 shows how the real estate cycle is lagged to the business cycle.
Moreover, Wheaton (1999) constructed a stock-flow model to study the office rental market and its cyclical behaviour of price development. He suggested that the supply of real estate is almost inelastic in the short term, meaning that the price for real estate is stabilized as demand meets supply in a market with constant supply.

However, this approximation is not applicable in the long run, where rent affects construction, and construction, in a longer perspective, affects supply. Consequently, the price elasticity of supply can be interpreted as the price elasticity of construction.

Moreover, the stock-flow model shows that the main factor causing cycles is the supply lag, which is caused by the time lag in construction of buildings. Wheaton (1999) questioned the idea of consistent volatility in the real estate cycle. He argues that longer construction lags lead to larger cycles. Differences between time lags of construction in different real estate sectors lead to differences in the risk of overproduction, meaning that fluctuations depend on the type of real estate, where retail and office properties are more prone to unpredictable changes in real estate cycles compared to residential properties.

Office and hotel markets are at greater risk of overproduction due to longer construction lags and higher uncertainty on the demand side. Residential market, on the other hand, experiences less uncertainty on the demand side, meaning that the risk of overproduction is lower. The construction lags for industrial real estate are shorter, which consequently means that the risk of overproduction is lower in this sector.
Furthermore, Wheaton (1999) argues that a high elasticity of demand reduces real estate cycles, while a high elasticity of supply is a precondition for cycles. The larger the differences between elasticity in supply and demand, the larger are the cycles. Additionally, Wheaton also suggests that myopic expectations leads to oscillations in the real estate cycle, while forward forecasting and rational expectations leads to smooth adjustment and stability. However, there are conflicts between rationality based and behaviour based theories.

3.2 Behaviour based theories

Brunes (2006) investigated three behavioural aspects; myopic behaviour, herd behaviour and overconfidence, and their influence on decisions made by investors in commercial buildings, up to the stage of the start of construction. A statistical part of the study was based on Tobin’s q-theory of investments, which states that investment decisions are based on the ratio of the price of an existing recently built property to the cost of production, also called Tobin’s q.

Brunes constructed a regression model to analyse construction volume and Tobin’s q with different time lags. Strong correlation between the volume of investment in a specific year, and Tobin's q in that specific year, could indicate myopic behaviour. Brunes found that the explanatory power of a model with a three-year lag on Tobin’s q was quite good. This could be interpreted as if investors in commercial buildings are characterized by myopic behaviour since the lag between the decision to build, and the completion of commercial buildings, is about three years on average.

However, Brunes hypothesis of herd behaviour and overconfidence among investors in commercial buildings did not get as strong support in his study as his hypothesis about myopic behaviour. (Brunes, 2006)

3.3 The q theory of housing investment

Tobin (1969) presented the idea that the “rate of investment – the speed at which investors wish to increase the capital stock – should be related, if to anything, to q, the value of capital relative to its replacement cost”. This idea is also known as the q theory, and the ratio, q, is known as Tobin’s q or the q ratio. Applied to the housing market, Tobin’s q can be expressed as the market price of an existing building divided by the total production cost for a similar building.

According to the q theory, it is profitable for housing suppliers to build when Tobin’s q is greater than one (q>1), since this indicates excess in demand on the housing market. On the other hand,
when Tobin’s q is less than one (q<1), there is an excess in supply, and housing production should stop. In this situation, it is cheaper to buy an existing house compared to a newly built one.

Several studies have found a clear statistical relationship between the size of Tobin’s q and the construction activity, both over time, and in different regions.

Jud and Winkler (2003) applied the q theory to the housing market. Using quarterly data for the United States, they present estimates of a q model of housing investment. Their results show a clear statistical relationship between the current and lagged values of the q ratio and housing investment. This suggests that housing developers seem to respond to the demands of housing consumers, building more when the prices of existing homes are high compared to the prices of new homes.

Furthermore, Berg and Berger (2006) investigated the existence of a stable relationship between Tobin’s q and the investment in owner-occupied houses, on the Swedish housing market between 1981 and 2003. Berg and Berger suggest that the major changes in economic policy in Sweden, in the early 1990s, resulted in a more market driven demand for housing investment. Their results indicate that there is a high degree of correlation between the housing investment and the q ratio for the last period of the sample, 1993-2003. However, a stable long-run relationship could not be detected between 1981 and 1992.

Moreover, Takala and Tuomala (1990) studied the housing investment in Finland between 1972 and 1987. Their findings suggest that the q ratio is a significant determinant for housing investment, after 1980.

3.4 Housing bubbles

According to Glaeser et al. (2008), housing price volatility are sometimes too high to be explained by fundamentals of the market. This suggests that over-optimism and irrational behaviour among buyers on the housing market could cause housing bubbles. Stiglitz (1990) defined a bubble as

“If the reason that the price is high today is only because investors believe that the selling price will be high tomorrow – when “fundamental factors” does not seem to justify such a price – then a bubble exists.”

Glaeser et al. (2008) argues that the volatility of housing prices differs across market with different supply elasticity. By examining housing prices, new construction, and supply elasticity during booms and busts, they showed that places with a more inelastic supply had much larger increases in prices and smaller increases in new construction during booms compared to more elastic places.
However, places with more elastic supply might suffer from larger welfare losses, as a result of a housing bubble, since there are more overbuilding in these places during a bubble.

3.5 The DiPasquale-Wheaton model

DiPasquale and Wheaton (1992), divided the real estate market into two interrelated markets: the market for real estate space and the market for real estate assets. Moreover, they presented a model that illustrates the important connection between these two markets, and how they are affected by macroeconomic factors and financial markets. The model demonstrates the effects, resulting from various exogenous shocks on rents, prices, construction and the stock of real estate. The model is presented in a four-quadrant diagram, which illustrates how changes in both the nation’s macro economy and its financial markets results in new equilibrium in the asset and property markets. (See figure 4)

![Image of the DiPasquale-Wheaton model](image)

*Figure 4. The DiPasquale-Wheaton model, illustrating the connection between the property market and the asset market. (DiPasquale and Wheaton, 1992)*

Moreover, DiPasquale and Wheaton (1992) argued that an increase in employment, the number of households, or household income would increase the demand for space, leading to rising rents and housing prices, which in turn leads to an increase in construction. Furthermore, the prices of real estate assets, in relation to its cost of construction, affect the supply of real estate. In the long run,
market prices and construction cost will be equal, while in the short term, these two may differ significantly because of the lags inherent in the construction process. (DiPasquale and Wheaton, 1992)
4 Methodology and research design

The following chapter aims to describe and justify the overall methodological approach for investigating and answer the research question; “Can housing construction in Stockholm be explained by fundamental economic factors, and if, to what extent?” A statistical analysis framework is presented as well as a description of how the thesis is conducted. Finally, the regression method is described together with a description of statistical tests performed on the regression model.

Due to the nature of the research problem, a deductive approach was applied. A deductive approach begins with the development of hypotheses based on existing theory, after which the hypotheses are tested by applying relevant methods. This either leads to the confirmation or rejection of the hypotheses. The general steps in studies using a deductive approach is presented in figure 5. (Saunders, et.al., 2016)

A quantitative method was used to answer the research question in this thesis. First, a literature review was conducted, after which the quantitative analysis was performed as a regression analysis.

4.1 Literature review as a theoretical basis for the research

First, to increase knowledge, and to achieve better understanding of the research subject, a literature review on previous studies on housing supply and demand determinants, as well as real estate cycles, was conducted. Mostly academic papers from economic journals were used in the literature review, but also government reports, and market reports published by well established companies, have been used to increase knowledge and to obtain information about Stockholm County and its housing market. The literature review serves as a base for choosing the variables used in the
regression analysis, as well as it investigates existing theory, which the hypotheses, formed in this thesis, are based on. Moreover, the literature review also serves as a complement to the regression model, giving the thesis a broader perspective. The literature review is presented in chapter 3.

4.2 Quantitative method

A quantitative analysis was carried out as a regression analysis, where data on housing construction in Stockholm County from 1992 to 2016 was analysed as the dependent variable. The regression analysis covers fundamental factors, serving as independent variables in the regression model. The following eight independent variables were examined: Repo rate, inflation, GDP, population, income, unemployment, housing prices, and construction cost.

The layout of the quantitative study in this thesis consists of the following steps: Firstly, the collection and transformation of data. Secondly, the modelling of a well fitted and reliable model. And finally, the interpretation and analysis of the results.

4.3 Regression analysis

Regression analysis is a collection of statistical techniques for investigating the relationship between variables. A regression model aims at modelling the effect of a given set of independent variables $x_1, x_2, ..., x_k$ on a variable $y$ of primary interest, called the dependent variable. (Fahrmeir et al., 2013)

4.3.1 Linear regression models

A commonly used type of regression model is the linear regression model, in which the dependent variable, $y$, is a linear function of the unknown regression coefficients $\beta_0, \beta_1, ..., \beta_k$, and the independent variables $x_1, x_2, ..., x_k$. The general multiple linear regression model can be written as

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon$$

Where $\beta_0$ is the intercept, and $\epsilon$ is the error term which contains unobserved factors, that is, factors other than the independent variables in the model, that affect the dependent variable. (Wooldridge, 2013)

Linear regression models are commonly used in empirical studies that investigate complex and unknown functional relationship between a dependent variable and some independent variables. (Myers et al., 2012) This thesis uses multiple linear regression models on time series data to analyse what fundamental economic factors that affect housing construction in Stockholm. Moreover, a
technique called Ordinary Least Squares (OLS) is used to estimate the coefficients in the regression model. The coefficients are obtained by minimizing the sum of the squared differences between a linear combination of the independent variables and the dependent variable.

4.3.2 Distributed lag models

Previous studies have shown that it takes time for an economy to fully react to changes in supply. The full reaction may appear after several time periods. When an independent variable is expected to affect the dependent variable after a time delay, the independent variable must be lagged in the regression model. This can be done in a distributed lag model, where one or more independent variables can affect the dependent variable with a time lag. (Wooldridge, 2013)

If $x_1$ is expected to affect $y$ with a lag of one period, while $x_2$ is expected to have an instant impact on $y$, the following simple lag model can be used

$$ y_t = \beta_0 + \beta_1 x_{1t-1} + \beta_2 x_{2t} + \epsilon_t $$

Moreover, if the regression model predicts the current value of a dependent variable $y$, based on both the current value, and different lagged values, of an independent variable $x$, the following distributed lag model is applicable

$$ y_t = \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \cdots + \beta_p x_{t-p} + \epsilon $$

However, lagged variables are associated with some problems, such as multicollinearity and an increased degree of freedom in the model. These problems can be checked using various tests presented in chapter 4.4. (Studenmund, 2014)

4.3.3 Time trend

It is common for economic time series to have a common tendency to grow over time. In order to draw causal inference using time series data, it is important to identify these time trends. Ignoring trending behaviour among variables in the regression model can lead to false conclusions that changes in one variable are caused by changes in another variable. However, a time trend variable can be included in the regression analysis to eliminate problems with spurious regression caused by trending variables. (Wooldridge, 2013)

Moreover, if the dependent and independent variable have different types of trends, e.g. one variable is trending upward and one downward, the inclusion of a time trend can increase the significance of the independent variable. (Wooldridge, 2013)
To account for the trending behaviour of the variables included in the regression analysis in this thesis, a time trend has been added to the regression model.

4.4 Statistical tests of the data and models

4.4.1 Stationarity

A stationary time series means that the variable has a constant mean and variance over time. In order to get a valuable result from the regression analysis, all time series in the model need to be stationary. However, it is common for time series to be non-stationary, and thus, show a random walk. The Dickey-Fuller test can be used to check for stationarity among all variables in the regression model. To deal with problems of non-stationarity, the time series can be transformed in several ways, e.g. through logarithmic transformation or first differentiation.

4.4.2 Correlation

Correlation is a statistic measurement that describes the degree of relationship between two variables. Correlation is described by a value between -1 and 1, where 1 describes perfect correlation, meaning that the variables follow each other to 100% and -1 describes perfect negative correlation, meaning that the variables move in opposite directions. A correlation coefficient of zero implies that there is no relationship between the variables, and thus, they move totally independent of each other.

4.4.3 Multicollinearity

Multicollinearity is a phenomenon in which there is a high correlation between two or more independent variables in a multiple linear regression model. Since multicollinearity usually leads to less reliable p-values for the independent variables together with wider confidence intervals, it can cause difficulties in distinguishing the individual effects of the correlated variables in the regression model.

To identify multicollinearity in a model, a variance inflation factor (VIF) analysis can be conducted. The variance inflation factor measures to what extent a variable is contributing to the standard error in the regression model, where a large VIF value indicates that issues with multicollinearity exist in the model. Usually, some degree of multicollinearity exists, and it is therefore important to identify whether it constitutes a problem in the regression model. After identifying variables with multicollinearity issues, correlated variables can be combined or removed to solve the problem. Furthermore, multicollinearity caused by common trends can be eliminated by transforming the time series using the first differences or the logarithms. (Studenmund, 2014)
4.4.4 **Autocorrelation**

Autocorrelation is the correlation between a given time series and a lagged version of the same time series. In other words, autocorrelation describes the degree of similarity between observations of the same variable over a period of time. If the observations are autocorrelated it means that future observations are affected by prior values, and, consequently, the observations are not random.

Furthermore, autocorrelation occurs most frequently in time series data sets. Since autocorrelation can cause bias it is important to test for autocorrelation among all variables in a regression model. The most commonly used test for autocorrelation is the Durbin-Watson d Test. Moreover, Durbin’s alternative test for autocorrelation can be used as a more general test to determining whether there is any evidence of autocorrelation. This test considers more than one lag, and it is therefore suitable for regression models that includes several time lags. (Studenmund, 2014)

4.4.5 **Heteroscedasticity**

Heteroscedasticity describes a situation in which the size of the error term differs across values of an independent variable. When the variance for the distribution of the error term is non-constant, the OLS regression tends to generate inaccurate estimates of the standard error of the coefficients, meaning that the standard errors are bias. Since the standard error plays a central role when conducting significance tests, bias can lead to incorrect conclusions about the significance of the coefficients in the regression model. The opposite of heteroscedasticity is homoscedasticity, which means that the error term has a constant variance. Figure 6 illustrates the difference between homo- and heteroscedastic variances. (Studenmund, 2014)

Moreover, the White test can be used to test for heteroscedasticity. Problems with heteroscedasticity can be solved by applying logarithmic transformation to the heteroskedastic variable. (Studenmund, 2014)

Figure 6. Illustration of homo- and heteroscedastic variances. a) funnel-shaped heteroscedastic variance, errors. b) homoscedastic variance, errors. (Fahrmeir et al., 2013)
5 Data collection and hypotheses

This chapter describes the data collection process, the data sources used and the transformation of the data. Furthermore, the dependent, and all the independent variables that will be included in the regression analysis are presented, together with their theoretical impact on housing construction as well as a derived hypothesis for each independent variable.

Berg and Berger (2006) argued that the changes in tax and housing policy in Sweden in the end of the 1980s and the beginning of the 1990s led to a more market driven housing demand. This means that determinants for housing investments might differ between the old and the new policy regime. According to Berg and Berger (2006) the full effect of the policy changes could be seen around 1992 and 1993. This forms the basis for the choice of time span for the data included in the regression in this thesis. All variables used in the regression analysis consist of time series data compiled on a quarterly basis for the period 1992-2016.

The data used in the regression analysis were collected from two different sources. Most data were retrieved from Statistics Sweden (SCB), but also the central bank of Sweden (Riksbanken) was used as a source for the data collection. All variables are presented in table 1, along with their data sources. Furthermore, nominal values are given for all variables containing monetary data. This means that the effect of the inflation on these variables is the same over time.

During the data collection, problems arose with finding some of the data on a quarterly basis. Population, income, and unemployment were only found on a yearly basis. Data on the annual percentage average of unemployment in Sweden were transformed into quarterly data by applying the same unemployment rate for quarter 1-4 for each year.

The yearly data on population and income were transformed into quarterly data by dividing the change in population and income, from one year to the next, by four. This transformation method is motivated by the fact that these variables are less volatile and the growth of these variables are approximately continuous. Furthermore, the consumer price index (CPI) was used as a measurement of inflation. Data on CPI was only found monthly. The monthly data was transformed into quarterly data by collecting the consumer price index corresponding to the first month in each quarter. Moreover, the income and unemployment data refer to Sweden, and not specifically to Stockholm County.
Table 1. Regression variables and their data sources

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing construction</td>
<td>SCB</td>
</tr>
<tr>
<td>Repo rate</td>
<td>Riksbanken</td>
</tr>
<tr>
<td>Inflation</td>
<td>SCB</td>
</tr>
<tr>
<td>GDP</td>
<td>SCB</td>
</tr>
<tr>
<td>Population</td>
<td>SCB</td>
</tr>
<tr>
<td>Income</td>
<td>SCB</td>
</tr>
<tr>
<td>Unemployment</td>
<td>SCB</td>
</tr>
<tr>
<td>Housing prices</td>
<td>SCB</td>
</tr>
<tr>
<td>Construction cost</td>
<td>SCB</td>
</tr>
</tbody>
</table>

5.1 Variables

5.1.1 The dependent variable

Housing construction
The dependent variable that is aimed to be explained in the regression model is the housing construction in Stockholm County. The housing construction data collected comprises the number of housing starts in newly constructed multi-dwelling buildings in Stockholm County during the period 1992-2016. The data have been retrieved from Statistics Sweden.

5.1.2 Independent variables and hypotheses
For each independent variable, a hypothesis is formulated, based on previous studies presented in the literature review. All hypotheses denote the effect on the dependant variable, housing construction, resulting from a positive change in the independent variable, when all other variables are held constant.

The null hypothesis states that the independent variable does not have an effect on the dependent variable, which can be written as

$$H_0: \beta_j = 0, \text{ where } j = 1, 2, 3, \ldots, 9$$

All hypotheses are summarized in table 2.
Repo rate
The repo rate is the overnight interest rate offered to banks by the central bank of Sweden, Riksbanken, and it is the primary instrument for adjusting inflation and economic growth. The central bank can reduce the repo rate to stimulate consumption in the country, and thereby increase inflation, or vice versa.

The repo rate was first introduced in Sweden in 1994. Before that, from December 1985 until May 1994, the marginal rate was the central bank of Sweden's policy rate. For the first nine data observations in the regression model in this thesis (1992 Q1 – 1994 Q1), the marginal rate is used as a proxy for the repo rate.

Furthermore, the repo rate in Sweden is forecasted to rise at a slow pace, but it is expected to remain at a low level over a foreseeable future. (Riksbanken, 2017) Mortgage rate follows the changes in the repo rate, and since low interest rate leads to decreased cost of mortgage capital for households, a low interest rate environment is expected to increase housing prices, which in turn is expected to lead to increased housing construction.

\[ H_1: \text{A positive change in repo rate yields a decrease in housing construction} \]

Inflation
Inflation is a sustained rise in the general level of prices of goods and services in an economy. With a rising price level, fewer goods and services can be bought with a unit of currency, meaning that inflation leads to a reduction in the purchasing power. A common measure of inflation is the inflation rate, which is the annual percentages change in a general price index. (Blanchard, 2000) The central bank of Sweden uses changes in the consumer price index, CPI, as a measurement of the inflation rate in Sweden. (Riksbanken, 2018)

Several studies have examined if investment in real estate could hedge against expected and unexpected inflation. The general result from these studies show that real estate provides a partial hedge against some components of inflation (Chaney and Hoesli, 2010). Moreover, Schwab (1982) investigated the relationship between inflation expectations and the demand for housing. According to Schwab, some consumer uses housing as a hedge against expected inflation. This means that a positive change in the inflation rate is expected to decrease the demand for housing, which in turn leads to a decrease in housing construction.

\[ H_2: \text{A positive change in inflation yields a decrease in housing construction} \]
**GDP**

The Gross Domestic Product, or GDP, is the value of all the goods and services produced in a country during a specific time period, usually one year. GDP is a way to measure the economic activity in a country and it is commonly used as an indicator of a country’s economic performance. GDP serves as a measurement of the size of a country’s economy, while GDP growth describes the performance of the country from year to year.

Demand and supply are fundamental factors that lead to the construction of GDP, which plays an important role in the business cycle. A growing economy usually leads to an increase in demand for goods and services as well as an increase in supply of these goods and services. Based on this, a positive GDP growth can be expected to increase the demand for housing, and thus also increase housing construction.

Moreover, Barras (1994) showed how the real estate cycle can be linked to the parallel business cycle, and how an economic upturn leads to increased property demand, which in turn leads to supply shortage, after which rents rise, eventually leading to a building boom.

Moreover, DiPasquale and Wheaton (1992) argued that economic growth increases all equilibrium variables, i.e. rent, price, construction, and stock, in the real estate market.

\[ H_3: \text{A positive change in GDP yields an increase in housing construction} \]

**Population**

Fanning (2005) argued that long term real estate cycles are partly affected by the national changes in population, where a growing population leads to an increase in demand for real estate.

Moreover, according to DiPasquale and Wheaton (1992), the demand for space depends, among other factors, on the number of households, which is closely correlated to the population. All else equal, an increase in the number of households will increase the demand for space, hence the housing construction increases, which eventually leads to an increase in the housing supply.

\[ H_4: \text{A positive change in population growth yields an increase in housing construction} \]

**Income**

Fanning (2005) suggested that long term real estate cycles are partly affected by the national changes in income, where a rising income lead to an increase in demand for real estate.

Moreover, DiPasquale and Wheaton (1992) argued that the household demand for space partly depends on household income, since the annual payments that a household can afford to pay are
primary determined by its level of income. An increase in household income is therefore expected to increase housing demand, leading to an increase in housing construction.

\( H_5: \) A positive change in income yields an increase in housing construction

**Unemployment**

Unemployment is the total number of people who are not employed but are actively searching for employment. Unemployment is usually measured as the unemployment rate, which is the number of people unemployed divided by all people currently in the labour force. (Blanchard, 2010)

DiPasquale and Wheaton (1992) argued that, among other variables, employment affects the demand for space. An increase in employment increases the demand for space, leading to an increase in rents. This in turn, leads to higher asset prices which generate an upturn in construction.

\( H_6: \) A positive change in unemployment yields a decrease in housing construction

**Housing prices**

Because of difficulties in finding data on housing prices for multi-dwelling buildings from 1992 to 2016, a real estate price index, which refers to one- and two-family houses is used as a proxy for housing prices for multi-dwelling buildings. A correlation test, and a comparison between the real estate price index and data on housing prices for multi-dwelling buildings during the period 2005-2016 (available from Statistic Sweden) showed a high degree of correlation and similarities between these two data sets.

DiPasquale and Wheaton (1992) argue that housing prices largely depends on supply and demand for housing. All else equal, an increase in demand for housing will raise housing prices, while an increase in housing supply will lower the prices.

Mayer and Somerville (2000) presented an empirical model of housing supply, which describes new housing construction as a function of changes in current and lagged housing prices. Their findings suggest that a 10% increase in real house prices leads to an 0.8% rise in housing supply.

\( H_7: \) A positive change in housing prices yields an increase in housing construction

**Construction cost**

To capture the effect of a change in construction cost in the regression model, an index called building price index (BPI) will be used. The building price index measures the development of prices for constructing new residential buildings, adjusted for quality and regional differences. The
The building price index consists of construction cost and connection fees, for example, electricity and district heating, as well as administrative costs, design costs, interest expenses, and value-added tax. The building price index used in this thesis is based on prices for multi-dwelling buildings.

According to the q theory, it is profitable for housing suppliers to build when Tobin’s q is greater than one, which means that the market price of an existing building is higher than the total production cost for a similar building. The q theory of housing investment implies that, all else equal, an increase in the construction cost would decrease construction.

\( H_8: \text{A positive change in construction cost yields a decrease in housing construction} \)

**Time trend**

To account for the trending behaviour of the variables in the regression model, a time trend has been added to the regression model. The independent variable shows a growing trend over time. This suggests that an increase in the time trend variable yields an increase in housing construction.

\( H_9: \text{A positive change in time yields an increase in housing construction} \)

**Table 2. Summary of hypotheses**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_1 ) Repo rate</td>
<td>-</td>
</tr>
<tr>
<td>( H_2 ) Inflation</td>
<td>-</td>
</tr>
<tr>
<td>( H_3 ) GDP</td>
<td>+</td>
</tr>
<tr>
<td>( H_4 ) Population</td>
<td>+</td>
</tr>
<tr>
<td>( H_5 ) Income</td>
<td>+</td>
</tr>
<tr>
<td>( H_6 ) Unemployment</td>
<td>-</td>
</tr>
<tr>
<td>( H_7 ) Housing prices</td>
<td>+</td>
</tr>
<tr>
<td>( H_8 ) Construction cost</td>
<td>-</td>
</tr>
<tr>
<td>( H_9 ) Time trend</td>
<td>+</td>
</tr>
</tbody>
</table>
6 Results

This chapter presents the results from the regression analysis. The regression outcomes will be presented followed by a summary of the hypothesis testing.

Figure 7 illustrates the changes in the dependent variable in the regression model. The figure shows the quarterly number of housing starts in newly constructed multi-dwelling buildings in Stockholm County between 1992 and 2016.

![Figure 7. Quarterly number of housing starts in newly constructed multi-dwelling buildings in Stockholm County, 1992 Q1 - 2016 Q4](image)

6.1 Regression outcomes

Table 3 presents a summary of the regression output. The full regression output is presented in Appendix I. All variables in the regression model have been log transformed and/or differentiated in order to retrieve stationarity among all time series. Tests for stationarity, using the Dickey-Fuller test have been conducted. The test results indicated stationarity among all variables, except population, which had a test statistic-value that was slightly higher than the 5% critical value. Moreover, to retrieve statistically significant estimates, some variables have been lagged up to seven periods, where one period equals one quarter. A correlation test has been conducted for all variables.
The test did not show any problems with correlation among any of the variables. The correlation test and the Dickey-Fuller tests are presented in Appendix II.

Furthermore, the regression model has been tested for multicollinearity, autocorrelation, and heteroscedasticity, using the variance inflation factor (VIF), Durbin’s alternative test, and White’s test, respectively. No problems with multicollinearity, autocorrelation or heteroscedasticity were encountered. However, since most economic variables are heteroscedastic by nature, the commando \textit{vce(robust)} has been applied to the regression model. This commando improves models that might suffer from heteroscedasticity, by making the standard errors robust. The commando improves the standard errors, but it does not affect the $\beta$ coefficients in the model. For detailed test statistics, see Appendix II.

Moreover, the model has an R-squared value of 0.73, which indicates that the model can explain 73% of the variation in housing construction in Stockholm County between 1992 and 2016.

\textit{Table 3. Regression results}

<table>
<thead>
<tr>
<th>Linear regression</th>
<th>Number of obs</th>
<th>R-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>91</td>
<td>0.7331</td>
</tr>
</tbody>
</table>

| Housing construction | Lag  | $\beta$ Coef. | Std. Err. | t | $P > |t|$ | Robust |
|----------------------|------|----------------|-----------|---|-------|--------|
| Repo rate            | Lag 7| -0.2792        | 0.1234    | -2.26 | 0.026** |
| Inflation            | Lag 7| 0.1521         | 0.0697    | 2.18 | 0.032** |
| GDP                  | Lag 0| 4.0898         | 1.0599    | 3.86 | 0.000***|
| Population           | Lag 1| -211.4862      | 51.9763   | -4.07 | 0.000***|
| Unemployment         | Lag 6| -1.1323        | 0.5182    | -2.19 | 0.032** |
| Housing prices       | Lag 2| 6.4807         | 1.9787    | 3.28 | 0.002***|
| Construction cost    | Lag 3| -0.9505        | 0.5484    | -1.73 | 0.087*   |
| Time trend           | -    | 0.0296         | 0.0022    | 13.41 | 0.000***|
| _cons                | -    | 5.8683         | 0.1623    | 36.15 | 0.000***|

\textit{Notes}: *, **, *** denote significance at 90%, 95%, and 99%, respectively.

The model presented in table 3 aims to explain the housing construction in Stockholm County between 1992 and 2016. P-values and $\beta$ coefficients are presented for all independent variables included in the model. Income was not statistically significant, and the variable was therefore excluded from the model.
The p-values are used to determine statistical significance in the hypothesis testing. If the p-value is less than 0.05, the null hypothesis can be rejected, meaning that the corresponding independent variable has a significant effect on the dependent variable to a 95% significance level.

The $\beta$ coefficient describes to what extent a change in a variable affect housing construction in Stockholm County. A positive value of the $\beta$ coefficient means that a positive change in the corresponding independent variable yields an increase in housing construction, while a negative value of the $\beta$ coefficient means that a positive change in the independent variable yields a decrease in housing construction.

Several variables were expected to have a lagged impact on housing construction. The model was constructed using different periods of lag on the independent variables.

The final model includes Repo Rate lagged seven periods, Housing Prices lagged two periods, Construction Cost lagged three periods, Population lagged one period, GDP without lag, Unemployment lagged six periods, and Inflation lagged seven periods. Also, a time trend, capturing trending behaviour of the model variables, was included in the regression model.

Furthermore, the dependent variable, Housing Construction, has been transformed using the natural logarithm. Moreover, in order to obtain stationarity, the Repo Rate, Housing Prices, Construction Cost, Population, and Unemployment have been transformed using the change in the natural logarithm, which can be interpreted as the growth rate of the variables. GDP has been transformed using the nature logarithm, and Inflation has been first differentiated.
6.2 Hypothesis testing

The hypothesis testing is presented in table 4, which includes a summary of the hypotheses together with the outcome from the regression model. Six out of nine hypotheses were correctly predicted to a significance level of 0.1 or above.

Table 4. Summary of the tested hypotheses with prediction and outcome

<table>
<thead>
<tr>
<th>Variable</th>
<th>Prediction</th>
<th>Outcome</th>
<th>Significance level</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₁ Repo rate</td>
<td>-</td>
<td>-</td>
<td>**</td>
</tr>
<tr>
<td>H₂ Inflation</td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>H₃ GDP</td>
<td>+</td>
<td>+</td>
<td>***</td>
</tr>
<tr>
<td>H₄ Population</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>H₅ Income</td>
<td>+</td>
<td>omitted</td>
<td></td>
</tr>
<tr>
<td>H₆ Unemployment</td>
<td>-</td>
<td>-</td>
<td>**</td>
</tr>
<tr>
<td>H₇ Housing prices</td>
<td>+</td>
<td>+</td>
<td>***</td>
</tr>
<tr>
<td>H₈ Construction cost</td>
<td>-</td>
<td>-</td>
<td>*</td>
</tr>
<tr>
<td>H₉ Time trend</td>
<td>+</td>
<td>+</td>
<td>***</td>
</tr>
</tbody>
</table>

Notes: *, **, *** correctly predicted and significant at 90%, 95%, and 99%, respectively.

Since the income variable is not significant at any acceptable level, there is not enough evidence to reject the null hypothesis, which states that there is no linear relationship between income and housing construction in Stockholm County. Therefore, the income variable is excluded from the model. Moreover, the hypotheses for Repo Rate, GDP, Unemployment, Housing Prices, Construction Cost, and the time trend were correctly predicted to a significant level of 0.1 or above, while the hypotheses for inflation and population was wrong.
7 Analysis

This chapter presents an analysis and discussion regarding the results and hypothesis testing presented in the previous chapter. All variables will be analysed and the discussions will reconnect to the literature review. Moreover, reliability and validity, as well as the limitations of the study, will be discussed in this chapter.

7.1 Analysis of variables

Six out of nine variables, examined in the regression analysis, match the predicted outcomes in the regression, as well as theory from previous studies presented in the literature review. However, there are some interesting and unexpected results that will be further discussed in this chapter.

Repo rate

A positive change in repo rate was expected to yield a decrease in housing construction. The hypothesis testing shows that the null hypothesis can be rejected to a 95% significance level. Interest rates, which are closely related to the repo rate, plays an important role for both the supply and demand side of the housing market. If the interest rates decrease, it becomes cheaper, and thus more profitable, for construction companies to build, which leads to an increase in supply. At the same time, lower interest rates lead to decreased cost of mortgage capital for households, which in turn leads to an increase in housing demand, as well as an increase in housing prices, which further enhance the willingness for housing developers to increase housing construction.

Inflation

A positive change in inflation was expected to yield a decrease in housing construction. However, the regression result showed a positive $\beta$ coefficient for the inflation variable. The result in the regression was unexpected since literature describes investment in real estate as a way to hedge against inflation. However, the result might be explained by the characteristics of booms and recessions. Booms are associated with periods of high inflation. Moreover, according to Barras (1994), economic booms are linked to building booms which in turn leads to increased supply. This linkage might explain the positive $\beta$ coefficient for the inflation variable in the regression outcome.
GDP
A positive change in GDP was expected to yield an increase in housing construction. The hypothesis testing shows that the null hypothesis can be rejected to a 99% significance level.

GDP growth describes the economic performance of the country. Moreover, supply and demand serves as important factors in the construction of GDP. An increase in supply and demand for goods and services can be expected to include an increase in the supply and demand for housing. During times when the economy is flourishing, people are better off, and can therefore afford more housing, which increase the demand for housing, leading to increased housing construction.

Population
A positive change in population was expected to yield an increase in housing construction. However, the regression result showed a negative $\beta$ coefficient for the population variable. This was a surprising and unexpected result. Literature suggest that the demand for space partly depends on the number of households, where an increase in the number of households would increase housing construction.

Several problems where encountered during the data collection and the data transformation process regarding the data on population in Stockholm County. Firstly, the data was only found on a yearly basis, which means that it had to be transformed into quarterly data. The transformation method used, where the change in population from one year to the next was divided by four to obtain quarterly data, generated a data set with a consistent population growth over each year. The actual population growth is probably more volatile. If quarterly data had been available, the result from the regression model might have been different and more accurate.

Moreover, the Dickey-Fuller test for stationarity showed a result indicating that the time series on population growth might experience problems with non-stationarity. The Dickey-Fuller test showed a test statistic-value slightly higher than the 5% critical value. The problem with non-stationarity might depend on the consistent population growth, resulting from the data transformation.

However, although this result was unexpected, the underlying reason for this result could also arise from several characteristics of the population growth. A more thoroughly analysis could have been conducted if the driving forces for the population growth was known. Housing construction is partly driven by the demand for housing, not by the need for housing. If the population growth mainly derives from an increase in the number of job opportunities in Stockholm, the demand for housing might rise. However, if the population growth is caused by increased immigration or increased childbirth in the region, the demand for housing may not rise to the same extent, due to a lower ability to pay for housing within these groups. Furthermore, the age distribution among the
population in a region can affect the demand for housing as well. A population with a high density of children might demand less housing since the population might consist of larger families living together. While a population consisting of a high density of young adults might increase the demand for housing. Moreover, if the population growth comes from an increase in childbirth in the region, the increase in the demand for housing might not show until after several years, and the time lag of up to three years, used in the regression model in this thesis, might not be enough to capture the effect of population growth.

**Income**

A positive change in income was expected to yield an increase in housing construction. However, the hypothesis could not be confirmed nor denied since the regression output showed an insignificant result for the income variable. Previous studies have suggested that the household demand for space partly depends on household income, since the annual payments that a household can afford to pay for housing are primary determined by its level of income. However, the insignificance of the income variable might be explained by inaccurate income data. The data was only found on a yearly basis, and it was therefore transformed into quarterly data by dividing the income growth from one year to the next by four, leading to a consistent income growth over each year. The actual income growth might be more volatile than the income data used in this regression, which might affect the result of the regression. However, in general, income is expected to grow over time, and a decline in income is rare. Therefore, the dataset is probably not too inaccurate. Moreover, the income data refers to Sweden, and not specifically to Stockholm County. While the percentage income growth can be expected to be approximately the same all over Sweden, this approximation might be one of the reasons to the insignificance of the income variable in the regression.

**Unemployment**

A positive change in unemployment was expected to yield a decrease in housing construction. The hypothesis testing shows that the null hypothesis can be rejected to a 95% significance level. DiPasquale and Wheaton (1992) argued that an increase in employment increases the demand for space, which in turn should lead to an increase in construction. A rise in employment rate indicates a situation where the financial position of the population increases, which could be expected to increase the demand for housing. Moreover, an employment is generally a requirement for an individual to be granted a loan at a bank. Consequently, an increase in employment could be expected to increase the number of people with the ability to buy a house or an apartment.
Housing prices
A positive change in housing prices was expected to yield an increase in housing construction. The hypothesis testing shows that the null hypothesis can be rejected to a 99% significance level. This result is in accordance with Mayer and Somerville (2000), who presented an empirical model of housing supply, which describes new housing construction as a function of changes in current and lagged housing prices. Their model suggests that an increase in real house prices leads to a rise in housing supply.

Construction cost
A positive change in construction cost was expected to yield a decrease in housing construction. The hypothesis testing shows that the null hypothesis can be rejected to a 90% significance level.

It is reasonable that, all else equal, a rise in construction cost decreases housing construction, since a rise in construction cost lowers the profitability for the construction companies. Moreover, this is in accordance with the q theory of housing investment, which states that it is profitable for housing suppliers to build when the market price of an existing building is higher than the total production cost for a similar building.

Time trend
A positive change in time was expected to yield an increase in housing construction. The hypothesis testing shows that the null hypothesis can be rejected to a 99% significance level.

The time trend is statistically significant, and its $\beta$ coefficient implies an approximate 3% increase in housing construction per year, on average. The R-squared value in the regression model presented in this thesis is much higher compared to a similar model where the time trend has been excluded. This might partly be explained by the fact that a time trend captures other factors that are trending over time, outside the variables included in the regression model.

7.2 Limitations

This study explores housing construction through eight different variables. Since there are many other factors that may affect housing construction, the result in this study might hence suffer from variable selection bias. However, a thorough literature study has been conducted, which serves as a base for the variable selection.

Moreover, some difficulties arose during the data collection process. Data on income and population were only found on a yearly basis. The transformation of the data into quarterly data might cause problem, since these time series are only approximations of the variables and not the actual values.
Furthermore, since the independent variables might affect housing construction with multiple time lags, the regression model could possibly be improved by including multiple time lags for each variable, meaning that each independent variable could affect the dependent variable with multiple time lags.

Moreover, one question to consider when constructing a regression model is whether one or more of the independent variables could be jointly determined with the dependent variable, usually through an equilibrium mechanism. For example, the dependent variable in this thesis, housing construction, can be explained by several variables. However, some of these variables might, in turn, depend on housing construction. This simultaneity problem could have been solved by applying a method for estimating a simultaneous equations model, which is a model in the form of a set of linear simultaneous equations. However, applying the method for estimating simultaneous equations models is beyond the scope of this thesis, and therefore housing construction is regarded as the only dependent variable in the regression model.
8 Conclusion

The following chapter presents a summary of the main findings in this study. Furthermore, ideas for further research within the research area are proposed.

8.1 Concluding remarks

The overall purpose of this thesis was to analyse what fundamental economic factors that affects housing construction in Stockholm. Furthermore, a multiple linear regression model was presented, which aims to describe to what extent fundamental economic factors affect housing construction in Stockholm County.

Moreover, the R-squared of 0.73, in the model, indicates that 73% of the variation in housing construction in Stockholm County can be explained by repo rate, inflation, GDP, population, unemployment, housing prices, construction cost, and a time trend. However, the regression model showed a low significance level for income, thus this variable was excluded from the model. Furthermore, a time trend was included in the regression model. The time trend showed statistically significance, and its β coefficient implies an approximate 3% increase in housing construction per year, on average.

There are several factors, not included in the regression model, that might affect housing construction. The municipal planning monopoly implies that municipalities serve as key actors for housing development. Their decisions might not be based on fundamental economic factors.

Moreover, also other policy regulations in Sweden might influence the development on the housing market, such as changes in taxes, and the introduction of government subsidies for housing construction. These factors are not included in the model.

8.2 Suggestions for further research

Future studies on housing construction in Stockholm could focus on qualitative factors, such as the impact of the Swedish planning system. Moreover, it would be interesting to examine the driving forces, for municipalities, to increase housing construction. This could be done by studying how municipalities work with the development of detail development plans within the municipalities.
9 References


Brunes, F. (2006) Overbuilding in office markets: Are behavioural aspects important? *Department of Infrastructure Building and Real Estate Economics, Royal Institute of Technology (KTH), Stockholm*


Reed, Richard and Wu, Hao (2010), Understanding property cycles in a residential market, *Property management*


Appendices

Appendix I. Regression model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing construction</td>
<td>Byggande</td>
</tr>
<tr>
<td>Repo rate</td>
<td>REPO</td>
</tr>
<tr>
<td>Inflation</td>
<td>INF</td>
</tr>
<tr>
<td>GDP</td>
<td>GDP</td>
</tr>
<tr>
<td>Population</td>
<td>POPG</td>
</tr>
<tr>
<td>Income</td>
<td>INC</td>
</tr>
<tr>
<td>Unemployment</td>
<td>UNEMP</td>
</tr>
<tr>
<td>Housing prices</td>
<td>HOP</td>
</tr>
<tr>
<td>Construction cost</td>
<td>BPI</td>
</tr>
<tr>
<td>Time trend</td>
<td>quarter</td>
</tr>
</tbody>
</table>

Regression output

Linear regression

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>lnByggande</td>
<td>Coef.</td>
</tr>
<tr>
<td></td>
<td>Std. Err.</td>
</tr>
<tr>
<td></td>
<td>t</td>
</tr>
<tr>
<td></td>
<td>P&gt;</td>
</tr>
<tr>
<td></td>
<td>[95% Conf. Interval]</td>
</tr>
</tbody>
</table>

-4.2791809  0.123413  -2.26  0.026  -0.5246897  -0.336732

Number of obs = 91
F( 8, 82) = 33.45
Prob > F = 0.0000
R-squared = 0.7331
Root MSE = 0.43995
Appendix II. Statistical testing

Test for Stationarity

Repo rate

. dffuller lr7REPO

Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Z(t) )</td>
<td>-5.192</td>
<td>-3.524</td>
<td>-2.898</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for \( Z(t) = 0.0000 \)

Housing prices

. dffuller lr2HOP

Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Z(t) )</td>
<td>-5.236</td>
<td>-3.516</td>
<td>-2.893</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for \( Z(t) = 0.0000 \)

Construction cost

. dffuller lr3BPI

Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Z(t) )</td>
<td>-15.150</td>
<td>-3.517</td>
<td>-2.894</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for \( Z(t) = 0.0000 \)
Population

. dfuller lr1POP

Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(t)</td>
<td>-1.169</td>
<td>-3.514</td>
<td>-2.892</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for Z(t) = 0.6868

GDP

. dfuller rGDP

Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(t)</td>
<td>-48.621</td>
<td>-3.513</td>
<td>-2.892</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for Z(t) = 0.0000

Unemployment

. dfuller lr6UNEMP

Dickey-Fuller test for unit root

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(t)</td>
<td>-9.490</td>
<td>-3.521</td>
<td>-2.896</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for Z(t) = 0.0000
Inflation

. dfuller d7INF

Dickey-Fuller test for unit root  Number of obs = 91

<table>
<thead>
<tr>
<th>Test Statistic</th>
<th>1% Critical Value</th>
<th>5% Critical Value</th>
<th>10% Critical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(t)</td>
<td>-14.805</td>
<td>-3.523</td>
<td>-2.997</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for Z(t) = 0.0000

Test for Correlation

<table>
<thead>
<tr>
<th>lnByggene</th>
<th>ln7REPO</th>
<th>ln2MOP</th>
<th>ln3BFI</th>
<th>ln1POPG</th>
<th>rGDP</th>
<th>ln6UNEMP</th>
<th>d7INF quarter</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnByggene</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln7REPO</td>
<td>-0.1780</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln2MOP</td>
<td>0.1330</td>
<td>-0.1704</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln3BFI</td>
<td>0.0256</td>
<td>0.0006</td>
<td>0.2777</td>
<td>1.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ln1POPG</td>
<td>0.3499</td>
<td>-0.0974</td>
<td>-0.0556</td>
<td>0.0015</td>
<td>1.0000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rGDP</td>
<td>0.1762</td>
<td>0.0328</td>
<td>0.1408</td>
<td>0.2060</td>
<td>-0.0053</td>
<td>1.0000</td>
<td></td>
</tr>
<tr>
<td>ln6UNEMP</td>
<td>-0.0067</td>
<td>-0.2440</td>
<td>0.1004</td>
<td>-0.1995</td>
<td>0.0477</td>
<td>0.0008</td>
<td>1.0000</td>
</tr>
<tr>
<td>d7INF</td>
<td>-0.1494</td>
<td>0.1259</td>
<td>-0.0954</td>
<td>-0.1751</td>
<td>-0.0192</td>
<td>-0.7356</td>
<td>0.1031</td>
</tr>
<tr>
<td>quarter</td>
<td>0.7732</td>
<td>-0.1263</td>
<td>-0.0777</td>
<td>0.0273</td>
<td>0.6463</td>
<td>-0.0235</td>
<td>-0.0631</td>
</tr>
</tbody>
</table>

Test for Multicollinearity

<table>
<thead>
<tr>
<th>Variable</th>
<th>VIF</th>
<th>1/VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>d7INF</td>
<td>2.42</td>
<td>0.412641</td>
</tr>
<tr>
<td>rGDP</td>
<td>2.40</td>
<td>0.416215</td>
</tr>
<tr>
<td>quarter</td>
<td>1.80</td>
<td>0.554066</td>
</tr>
<tr>
<td>ln1POPG</td>
<td>1.75</td>
<td>0.572010</td>
</tr>
<tr>
<td>ln6UNEMP</td>
<td>1.20</td>
<td>0.833709</td>
</tr>
<tr>
<td>ln7REPO</td>
<td>1.20</td>
<td>0.834087</td>
</tr>
<tr>
<td>ln3BFI</td>
<td>1.19</td>
<td>0.841578</td>
</tr>
<tr>
<td>ln2MOP</td>
<td>1.16</td>
<td>0.862001</td>
</tr>
</tbody>
</table>

Mean VIF | 1.64
Test for Autocorrelation

Durbin-Watson d-statistic( 9, 91) = 1.849586

Durbin's alternative test for autocorrelation

<table>
<thead>
<tr>
<th>lags(p)</th>
<th>chi2</th>
<th>df</th>
<th>Prob &gt; chi2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.415</td>
<td>1</td>
<td>0.5196</td>
</tr>
<tr>
<td>2</td>
<td>1.272</td>
<td>2</td>
<td>0.5293</td>
</tr>
<tr>
<td>3</td>
<td>1.257</td>
<td>3</td>
<td>0.7395</td>
</tr>
<tr>
<td>6</td>
<td>3.479</td>
<td>6</td>
<td>0.7469</td>
</tr>
<tr>
<td>7</td>
<td>4.669</td>
<td>7</td>
<td>0.7603</td>
</tr>
</tbody>
</table>

H0: no serial correlation

Test for Heteroscedasticity

White's test for H0: homoskedasticity
against Ha: unrestricted heteroskedasticity

$\text{chi}^2(44) = 37.42$
Prob > chi2 = 0.7478

Cameron & Trivedi's decomposition of IM-test

<table>
<thead>
<tr>
<th>Source</th>
<th>chi2</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heteroskedasticity</td>
<td>37.42</td>
<td>44</td>
<td>0.7478</td>
</tr>
<tr>
<td>Skewness</td>
<td>5.58</td>
<td>8</td>
<td>0.6947</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>1.54</td>
<td>1</td>
<td>0.2150</td>
</tr>
<tr>
<td>Total</td>
<td>44.53</td>
<td>53</td>
<td>0.7896</td>
</tr>
</tbody>
</table>