NHH

Norwegian School of Economics
Bergen, Fall 2022

# House Prices - On a Kite, or in Free Flight?

*An empirical analysis of fundamental price in the Norwegian housing market, and its interaction with actual price.*

**Thomas Blekastad Davis and Ole Hallvard Toven Dæhli**

**Supervisor: Fred Schroyen**

Master thesis, Economics and Business Administration

Majors: Business Analytics & Financial Economics

NORWEGIAN SCHOOL OF ECONOMICS

# Acknowledgements

We want to express our sincere gratitude towards our supervisor Fred Schroyen who has contributed valuable and professional feedback in the creation of this thesis. His knowledge and helpfulness have proved invaluable. In writing this thesis, we have improved our knowledge about the Norwegian housing market, its fundamentals and statistical methods. Working with the thesis has been both engaging, educational and challenging as we have implemented what we have learned throughout our studies.

We would like to express our gratitude to Pål Boug at Statistics Norway for his assistance in obtaining data for this thesis. We also want to extend our thanks to our families and friends for their love and support throughout this journey. Without their help and encouragement, this achievement would not have been possible. Finally, we want to thank each other for our enjoyable collaboration and excellent teamwork this fall.

Norwegian School of Economics

Bergen, December 2022

Thomas Blekastad Davis

Ole Hallvard Toven Dæhli

# Abstract

In the past few decades, there has been a considerable increase in real house prices in Norway. It is therefore important to understand what this increase is based upon. In this study, we analyse the relationship between price and fundamental factors of the Norwegian housing market in the period of 1993 to 2022. Applying the methods of Bergman and Sørensen (2021), we estimate a five variable VAR model and compute a fundamental price. We find that prices in the Norwegian housing market have generally been aligned with what fundamentals would suggest across the analysed period. Additionally, we analyse the interaction between fundamental and actual price through a VECM, where we analyse whether there is cointegration between the prices, and if the interaction is consistent with the theory of fundamental valuation. We find that there is a cointegrated relationship, and that we cannot reject the null hypothesis that the relationship is in line with theory. These results are consistent when ending the sample from 2013 to 2022.

***Keywords*** – Norwegian Housing Market, Fundamental Valuation, VAR & VECM

# Contents

# List of Figures

# List of Tables

# 1 Introduction

## 1.1 Motivation and Purpose

As far as investments go, virtually nothing has been considered as "sure a thing" in Norway these last few decades as owning your own home. Indeed, according to Statistics Norway the average Norwegian home has increased in price by more than 200 % in real terms since the start of this millennium (SSB, 2022a,b).[1] This has been no small benefit to a considerable proportion of the population, as Norway has one of the higher rates of home ownership in the world, at more than 80% in 2020 (Eurostat, 2022). This historical development has provided a cushion of wealth creation that the average Norwegian household has been able to rely on. However, such a seemingly stable return on a specific commodity begs the question, how long can it last? Logically, the relative value of housing will surely not keep growing forever. In fact, there is empirical evidence that this has not always been the case. Eitrheim et al. (2004) find that real house prices in Norway did not reach their 1899 levels again until the mid-1980s. With this in mind, one might wonder why prices have been growing so steadily and whether there is any basis for the price level. In other words, is the price correct?

The question of whether house prices are accurate has been asked many times over the years by academics, policymakers, and consumers alike. Compared with other commodities, the scrutiny on house prices is particularly high, as a housing purchase is likely the largest investment most households will make. Therefore, many different theories and methods have been introduced over the years to describe, explain and evaluate the level of house prices. In this respect, one of the more established theories has been the theory of fundamental valuation (Campbell and Shiller, 1988a,b). From this theory, it is assumed that a *fundamental valuation* can be ascertained from underlying *fundamental factors*, which can explain the dividend of the asset and rational expectations about the future. Various methodologies have been constructed with this theoretical framework, and a range of fundamental prices are estimated across countries and time periods. One paper that has recently estimated a fundamental valuation of the housing market is titled "*The interaction of actual and fundamental house prices: A general model with an application*

[1] Real-term growth calculated as nominal growth subtracted for CPI development.

*to Sweden*", and is written by Michael Bergman and Peter Sørensen (2021). In it, they estimate a fundamental valuation of the Swedish housing market based on a *Supply and Demand model*. This model has its basis in the methods introduced by Campbell and Shiller (1988a) but extends their work in several ways.

With this thesis, we aim to provide new perspectives on the Norwegian housing market by estimating a fundamental house price and comparing it with the real price observed in the market. To do this, we will largely follow the methodology of Bergman and Sørensen. Although fundamental valuations of the Norwegian housing market have been conducted in the past, to our knowledge, this specific methodology has not been applied to Norwegian data. Therefore, we believe that testing the approach in Norway can provide new understanding and produce valuable insights.

## 1.2 Research Questions

In order to achieve this, we need to clarify our overall objectives. With this thesis, we aim to answer two specific research questions. As we have stated, we want to find what price level the fundamental factors imply that the price should be. Therefore, our first research question is as follows:

i: "*To what degree are Norwegian house prices in line with what fundamental factors would suggest?*"

When the fundamental price in the housing market is determined, it would also be interesting to ascertain whether there is any consistent relationship between the estimated price and the actual housing price. Furthermore, if there is a relation, do the prices interact in the way we expect from the theory of fundamental valuation? One assumption from theory is that the gap between the prices is stationary, i.e., that in the long run, prices are aligned. A second aspect is that a fundamental price should not be affected by a gap between itself and the real price. The implication is that only the actual price should react to a gap. With this in mind, we define our second research question:

ii: *" Is there a long-term relation between the fundamental and the actual house price, and if so, is it in line with what the theory of fundamental valuation would predict?"*

## 1.3 Main Findings

In this thesis, we estimate a fundamental price of the Norwegian housing market based on the methodology introduced by Bergman and Sørensen. In the period of Q2 1993 - Q2 2022, we find that the Norwegian housing market is largely in accordance with its fundamentals. Still, we also find that from 2016 until 2020, there was a period where the fundamental and actual prices moved in different directions, resulting in a significant gap between them. However, this gap seems to be closing towards the end of the sample.

Considering our second research question, we find that there is a stable, long-term relationship between fundamental and actual prices in the Norwegian housing market. Furthermore, we find that the estimated relationship is valid even when restricting the relationship such that it is in line with the theory of fundamental valuation. These findings are persistent when ending the sample from 2013 to 2022.

## 1.4 Structure of Thesis

In order to answer the research questions clearly and coherently, this thesis will have the following structure.

The next section introduces the relevant literature and theory on fundamental valuation. This is followed by a discussion of key theoretical concepts and models. The methodology for this thesis is then presented, followed by an overview of the data that will be used in the empirical analysis. Thereafter, we present the analysis, where we attempt to answer the two research questions. Finally, we discuss the results of the analysis and arrive at an overall conclusion for our thesis.

# 2 Literature Review

When considering the housing market, there is a vast range of published scientific papers and econometric concepts. As a result, the potential scope of relevant literature for our thesis is considerable. Therefore, in order to be expedient in our literature review, we will focus on the literature most pertinent to our thesis, i.e., a fundamental evaluation of house prices in Norway. With this in mind, we begin by introducing the theory of fundamental value and its use in the housing market, the findings of Bergman and Sørensen, as well as previous research on fundamental house prices in Norway.

## 2.1 The Housing Market and Fundamental Valuation

The concept of determining a fundamental value of an asset was introduced in 1988 by Campbell and Shiller with their dividend discount model (1988b; 1988a). They established a model for asset valuation, where the valuation is equal to the present value of the expected future cash flows generated by the asset, e.g., dividends. The agents in the model are assumed to be rational and form expectations through linear dynamics of fundamental factors. Here, fundamental factors are interpreted as relevant macroeconomic and microeconomic variables which should influence the expectations. In order to estimate the formation of expectations, they apply a vector autoregression (VAR) model to the variables. We will explain the VAR model in detail in our theoretical concepts section.

The dividend discount model was first introduced to model the behaviour of financial assets such as stocks. However, the methodology can be applied to the housing market as well. Where a liquid financial asset like a stock can provide a cash flow through dividends, owning (and using) a housing unit provides utility to the user. This utility can, thus, be compared to an agent's willingness to pay for housing services for a period, making it possible to derive a fundamental theoretical price for the housing market.

Following the methodology of Campbell and Shiller (1988a), multiple studies have estimated fundamental house prices across a variety of different markets, such as Hott and Monnin (2008), Campbell et al. (2009), and Ambrose et al. (2013). For example, Ambrose estimated the long-run relationship between rents and prices to determine fundamentals on 355 years of data on housing in Amsterdam. The subject of fundamental valuation is compelling because if fundamental factors can be used to determine a *fair value* of the housing market, then deviations away from this value can be scrutinised by markets and policymakers to, e.g., avoid housing bubbles. In fact, one can describe a housing bubble as a period where actual prices are decoupled from underlying fundamentals, reaching price levels considerably higher than what the fundamentals would justify (Stiglitz, 1990).

## 2.2 Bergman & Sørensen and a Fundamental House Price in Sweden

Michael Bergman and Peter Sørensen have published one such estimation of fundamental price in their paper, "The interaction of actual and fundamental house prices: A general model with an application to Sweden" (2021). As mentioned in the introduction, we aim to apply their methods to Norwegian data. Therefore we choose to devote some time in this section to explain their methods, motives, and results.

Their paper estimates the relationship between fundamental and actual housing prices in Sweden. In order to do so, they follow the procedure laid out by Hott and Monnin (2008) and base their analysis on the development of five variables; the actual house price, disposable income, housing stock, rent price, and the user cost of housing. In line with the procedure laid out by Campbell and Shiller, they then apply a vector autoregression model (VAR) in order to compute the coefficients needed to calculate the expectations of agents in their model of fundamental price. Moreover, they analyse the interactions between fundamental and actual prices and test the hypothesis that the actual price will tend to be drawn towards the level of the fundamental housing price. In contrast, the fundamental house price will be unaffected by changes in the actual price. Here they complement earlier research, such as Malpezzi's error correction model (1999), by creating a unified vector error correction model (VECM).

Through their approach, Bergman and Sørensen seek to answer questions such as whether prices converge towards fundamental price estimations and the speed of such a convergence. As part of this analysis, they identify the impact of shocks on both fundamental and actual housing prices and analyse how long it would take for prices to return to equilibrium. Finally, they also investigate the impact of feeding in an exogenous timeline of input variables on the fundamental price, simulating policy changes. This allows them to infer what effects such policy changes might plausibly cause on the fundamental valuation of the Swedish housing market.

Bergman and Sørensen estimate the fundamental price in the Swedish housing market from a quarterly data set spanning from 1986 to 2019. They find their time series to contain a structural break, diminishing the sample's statistical validity to perform the necessary analysis. The identified break is the Swedish banking crisis at the beginning of the 1990s. Therefore, the authors implement their approach on a subset with the first observation after the structural break. They estimate their fundamental price and find that the Swedish housing market is overpriced at the end of their sample, Q4 2019.

When analysing the interaction between the fundamental and actual price, Bergman and Sørensen find that the fundamental price has an anchor effect on the actual house price, i.e., that the actual price is drawn towards the level of the fundamental value. These results were in accordance with their expectations. This corroborates the earlier research of Hott and Monnin (2008), who, when examining data on the U.S., United Kingdom, Japan, and others from the 1970s to 2005, find that actual house prices often deviate from their fundamentals. Moreover, they find that these deviations can last for extended periods. They also find that the prices tend to return to equilibrium with fundamentals in the long run. When analysing the impact of shocks on either price, Bergman and Sørensen find that the subsequent gap has a half-life of several years, indicating a slow reversion to the mean.

## 2.3 Anundsen and a Fundamental House Price in Norway

In research published in 2019 and 2021, André Anundsen (2019, 2021) has estimated the fundamental value of the Norwegian housing market. The research aimed to determine whether there have been systematic overvaluations in the Nordic housing markets from 2000 to 2019. He applies various methods, including creating a VAR model where fundamental variables like real interest rate, housing stock, and disposable income are used to estimate the fundamental house price in Norway, Sweden, Denmark, and Finland.

Anundsen also uses the Johansen trace test to test for co-integration between the fundamental variables and the actual house price (2019). This is comparable to Bergman and Sørensen (2021), with the difference being that they analyse the estimated fundamental price for co-integration with the actual price and not the variables used in its estimation.

The research provides insightful results. He finds that the Norwegian housing market is priced at a discount from 2010 to 2016. After that, he finds the market is overpriced until the end of his sample, with a nine percentage overvaluation at the end of 2019. For Sweden, Anundsen finds that the Swedish housing market has been overvalued since 2014. At the end of 2019, this gap is estimated to be around seven per cent. Analysing the evolution of the fundamental price over the sample period, Anundsen finds that most of the increase in the fundamental valuation in the Nordic housing markets can be explained through a considerable increase in disposable income (Anundsen, 2021).

## 2.4 The Application of the Papers in this Thesis

Both Anundsen and Bergman and Sørensen investigate a fundamental valuation of Scandinavian housing markets for approximately the same time frame as our analysis will investigate. Moreover, the methodology applied in these papers broadly follows the same principles as ours, developing a VAR model to estimate the fundamental price and analysing the co-integration between the fundamental price and the actual price. We note that Anundsen uses the same methodology for Swedish and Norwegian data. Consequently, it is possible to compare Anundsen's results to the publications of Bergman and Sørensen, as well as the results of this thesis. This facilitates constructive discussions regarding results, assumptions, and overall methodology.

However, there are some differences between Anundsen and what this thesis will seek to do. Anundsen estimates a fundamental price based on a VAR model with three variables. In contrast, this paper follows Bergman and Sørensen (2021) in estimating the VAR model on the following five variables; the actual price, user cost, housing stock, disposable income, and the rent price of housing. Furthermore, the mathematical method used by Bergman and Sørensen to compute the fundamental price follows Hott and Monnin (2008), which is also the choice for this thesis. This is not the case with Anundsen, who applies a different methodological approach. Anundsen also bases his analysis on the assumption that the fundamental and actual house price is in equilibrium at the start of the sample in 2000. We do not follow this assumption.

These differences are likely to lead to somewhat different conclusions than what Anundsen found in his study, despite the otherwise notable similarities in methodology and data. Therefore, we see our research as complementary to Anundsen. Moreover, Anundsen's primary focus is on detecting housing bubbles. Instead, this thesis follows the procedure in Bergman and Sørensen with a more in-depth analysis of the nature of the relationship between the actual and the fundamental house price. Therefore, by going beyond the scope of Anundsen, this thesis can provide new insights into the Norwegian housing market while also providing more understanding of Anundsen's research by estimating the data with comparable methods.

# 3 Theoretical Concepts

As mentioned in the introduction, this thesis has two primary objectives. First, to estimate a fundamental price on the Norwegian housing market, and second, to analyse the relationship between this estimate and the actual price over time. This endeavour will demand the application of various methods, tests, and specific terminology with which the reader may or may not already be familiar. Therefore, we see it as apt to introduce some of these core concepts we will apply later in the analysis.

## 3.1 The Concept of Time Series Analysis

First and foremost, this thesis is an exercise in time series analysis, as we have observations of macroeconomic variables on a quarterly basis. As a result, it is helpful to understand what this means. Time series analysis is a branch of statistics and econometrics that focuses on the study of data collected over time (Lütkepohl, 2007). Time series data are typically collected at regular intervals, such as daily, monthly, or annually, and are most often used to study trends, patterns, and other changes in observed levels.

Time series can be broken down into sub-components. These sub-components are the trend, seasonal, and irregular components (ABS, 2022). The trend component indicates the long-term direction of the time series, the seasonal effects encapsulate any recurring variations in the data, while the irregular component is what is left, i.e., unsystematic movement and short-term fluctuations. This unsystematic movement can also be characterized as the stochastic component of a time series. Depending on the unit roots of the data, this stochastic component can either be stationary over time or include a random walk process (Pfaff, 2008).[2]

In order to conduct predictive analysis on time series with auto-regressive and moving average modes, the data is required to be stationary. A time series being stationary refers to it not changing its statistical properties over time (Palma, 2016), i.e., it exhibits a constant mean, variance, covariance and autocorrelation across time. Taken in the context of the components described above, this implies no trend or seasonal component.

[2]Unit roots are stochastic processes, where *unit root* = 1 means that there is a random walk process, where, e.g., yesterday's value is the best predictor of today's value. See chapters 1 and 3 of "Analysis of Integrated and Cointegrated Time Series with R" by Bernard Pfaff (2008) for further information.

However, often this is not the case. For example, macroeconomic data are usually not stationary as it describes the real world, where there are changes to, e.g., population levels, economic growth and investment levels, both over time and between seasons. Additionally, it is sometimes the case that the movement of macroeconomic variables includes a non-stationary stochastic component, i.e., a random walk (Pfaff, 2008). This increases the risk of spurious regression, i.e., finding statistically significant results between unrelated data (Hill et al., 2018).

As a result, it is common to transform the data prior to conducting analysis. A common transformation is taking the log of the variable(s), which transforms the data such that we have percentage changes instead of unit changes between observations. Another transformation is taking the difference of the variables. This is done when there is integration in the data. We can then identify an order of integration, $I(d)$, where $d$ denotes the number of differences that needs to be taken in order to produce a stationary time series. In this context, a time series with an integrated order $I(0)$ is stationary (Pfaff, 2008). In the case of time series with a random walk component, this can be shown to be a $I(1)$ time series, where taking the first difference will return a stationary time series (Lütkepohl, 2007).

## 3.2 Vector Autoregression Model (VAR)

In this thesis, we follow Bergman and Sørensen in using a vector autoregression model (VAR) in order to arrive at an estimate of a fundamental price. We have also seen that Campbell and Shiller utilised a VAR model back in 1988 to compute a fundamental valuation of company shares. However, exactly what a VAR model is and how it works is not necessarily instantly apparent to most people. Therefore, a section on theoretical concepts calls for an introduction to this specific type of model.

Christopher A. Sims introduced the concept of vector autoregression models (VAR) in his 1980 "Macroeconomics and Reality" as a tool to capture existing relationships between multiple variables that are allowed to change over time (Sims, 1980). The VAR($p$) model is defined such that it allows for $k$ number of equations and variables, where a linear regression estimates each variable based on the $p$ number of past values of itself and the other $k-1$ variables, and an error term, $\epsilon_t$ (Stock and Watson, 2001). As such, it is a multivariate model. This differentiates it from univariate autoregression models, which estimate a variable as a function of its own lagged values exclusively.

Sims (1980) presents three separate VAR models, the structured VAR model (SVAR), the recursive VAR model, and the reduced-form VAR model. In this thesis, we will focus on the reduced-form VAR model. Instead of attempting to identify the structural parameters that underlie the equations, as for example the SVAR model does, the reduced-form VAR model uses historical data to estimate the relationships between variables. Additionally, the error term of the reduced-form VAR model is distinguishable from the other VAR models by its serially uncorrelated error terms. In its general form, a reduced-form VAR($p$) model can be described in the following mathematical notation:

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \epsilon_t \tag{3.1}$$

Here, $y_t$ is a ($k$ $x$ $1$) vector and is the predicted value of the $k$ number of variables. The variable $c$ is defined as a ($k$ $x$ $1$)-vector of constants and is the model's intercepts. Meanwhile, $y_{t-i}$ is the $i$ number of ($k$ $x$ $1$)-vectors containing the prior observations of the variables, up to the lag length $p$.[3] $A_i$ is a ($k$ $x$ $k$)-matrix containing the estimated regression coefficients of the variables in a given lag length, while $\epsilon_t$ is the serially uncorrelated error term.

### 3.2.1 Checking the Residuals of a VAR model

When estimating a VAR model and deciding on the parameters, it is beneficial to conduct some statistical tests on the residuals the model will return, i.e., the error term. If a model has significant issues with the residuals, then this can draw into question the level of inference one can make from a model's estimations. Common aspects to test for include autocorrelation, heteroscedasticity, and normality of the residuals. In the following paragraphs, we will briefly explain these three aspects and the specific tests we will employ in our case to test for them.

[3] We generally define $i$ as being an expression for the values (1, ..., $p$ )

#### 3.2.1.1 Autocorrelation and the Breusch-Godfrey Test

Autocorrelation is when an observation can be seen as a function of lagged observations of the same variable (Palma, 2016). A variable can have positive and negative autocorrelation, which signifies that the value of the variable is positively or negatively correlated with the former observation(s). There is no autocorrelation if the past observations cannot predict the level or trend of the variable. Autocorrelation is an issue in the reduced-form VAR model and in time series generally, as it violates the assumption of serially uncorrelated error terms. Implicitly, it implies that better-specified models are available, as there is some trend or seasonal effect that the specified model cannot capture. In order to test for autocorrelation in a multivariate VAR model, we can employ the multivariate Breusch-Godfrey test for serial correlation (Breusch, 1978). The statistical procedure tests whether autocorrelation is present in the data and returns a p-value, with the null hypothesis that there is no autocorrelation in the error terms.

#### 3.2.1.2 Heteroscedasticity and the ARCH Test

Heteroscedasticity refers to situations where the variance of a time series's residuals fluctuates (Palma, 2016). More specifically, a time series is said to be heteroscedastic if there is a systematic change in how the residuals are spread out throughout the range of the measured values. A non-constant variance complicates analysis as this violates the assumptions of ordinary least squares (OLS) analysis. It also implies that the model varies in its predictive accuracy over time. This, in effect, means that there are some changing effects that the model is failing to estimate. Heteroscedasticity in a multivariate VAR model can be tested through an autoregressive conditional heteroscedasticity (ARCH) model. We will specifically use the ARCH-LM test by R.F. Engle, which describes the variance of the relevant error term as a function of actual sizes from the error terms of previous periods (Engle, 1982).[4] The test will return a p-value of the null hypothesis that there is no heteroscedasticity in the residuals.

[4] The autoregressive conditional heteroscedasticity - Lagrange multiplier (ARCH-LM) test was initially intended for univariate tests only, but multivariate versions have been developed. We will compute this extension of the test through the statistical software package R.

#### 3.2.1.3 Normality and the Jarque-Bera Test

Concluding the standard residual tests, normality refers to the distribution of the residuals. Specifically, it is a test to check whether they are normally distributed, as is the base assumption. The central limit theory states that when the sample increases, it is possible to infer that the data will be normally distributed (Fischer, 2011). Significant deviations from a normal distribution will impact the confidence one can have in many statistical tests' ability to discern the true difference in the data. In the case of a multivariate VAR model, a Jarque-Bera test can be applied to test for normality in the data. This is a goodness-of-fit test of whether the data has kurtosis and skewness corresponding to a normal distribution. A test statistic far from zero indicates that the data is not normally distributed (Jarque and Bera, 1980). The null hypothesis with this test is that the residuals of the VAR model are normally distributed.

### 3.2.2 Granger Causality: The Relative Importance of Variables

We also want to assess the VAR model estimates in our analysis. One applicable test in order to derive such insight is a test for Granger Causality. This test determines to what degree one variable is important in forecasting future values of other variables in the VAR model. Functionally, it works by conducting $t$-tests and $F$-tests on the lagged values of each variable and ascertaining the likeliness of that variable explaining future levels of the other variables.[5] As not to confuse Granger causality with actual causality, variables are said to "Granger-cause" other variables (Granger, 1969).

[5] $t$-tests and F-tests are statistical tests used in order to test for similarity of two populations mean and variance respectively. Varieties of the tests exist for different needs (Enders, 2008).

### 3.2.3 Cointegration and the Johansen Tests

When estimating the VAR model, we will also want to test for cointegration. Cointegration refers to the existence of a long-term equilibrium relationship between two or more time series (Pfaff, 2008). This equilibrium can be a difference of zero or some other stable ratio. Specifically, cointegration refers to the time series being non-stationary, i.e., of integration order $I(1)$, and there being some form of a linear combination of the variables that is stationary, i.e., $I(0)$ (Lütkepohl, 2007). In a multivariate system, there is a max of $k$ cointegration relationships, denoted as $r$. A rank of $r = k$ implies that the time series are already stationary with $I(0)$. We note that a cointegration relationship between two variables represents an exception to the general rule of using stationary data when conducting time series analysis. As a result, regression will not increase the risk of spurious regression (Hill et al., 2018).

Various tests allow for short-term deviations between the time series as long as the long-term relationship exists. One way to test for cointegration between the variables is the Johansen test (Johansen, 1988, 1991). The test evaluates the validity of a cointegration relationship through an approach of maximum likelihood estimates. The Johansen trace test is useful in the case of a multivariate VAR model, as it allows for several cointegration relationships. Moreover, it allows for short-term deviations as long as a long-term relationship between the time series exists Johansen (1991).

Describing the test in further detail, the null hypothesis of the Johansen trace test is that the actual number of cointegration vectors ($r^*$) are such that $r^* \leq r$, for $r = (0, 1, ..., (k-1))$. The alternative hypothesis for each evaluated rank is that $r^* > r$. Starting at a hypothesis of $r = 0$, the test is repeated until one cannot reject the null hypothesis. There is also a second version of the Johansen test, the Johansen eigen test. The tests are quite similar, with the main difference being that the alternative hypothesis is instead $r = r^* + 1$. Like the trace test, this test is evaluated on increasing ranks until the null hypothesis is not rejected (Hänninen, 1998).

## 3.3 Vector Error Correction Model (VECM)

When estimating a VAR model, one assumes that the time series in the model are stationary or that the variables are cointegrated with a full rank, i.e., $r = k$ and thus $I(0)$ (Hill et al., 2018). However, sometimes we have some cointegration between the variables in a multivariate system, but not full rank. Alternatively, we might be interested in further analysing the cointegration relationship between variables. A vector error correction model (VECM) can be applied in both cases. A VECM is, in its purest definition, a VAR model which is stationary in first differences, $I(1)$ (Lütkepohl, 2007). In other words, it takes the time series and estimates a VAR model with the lag order $p - 1$, where $p$ is the estimated lag order of the VAR model on the original data.

Crucially, the VECM can account for the number of cointegration relationships between the variables. Thus it is an option when a VAR model is not of full rank, and to see if the VECM returns similar results. Furthermore, a VECM can test hypotheses regarding the cointegration relationships between variables. As its name applies, it is an error correction model. The error correction here refers to the model's assumption that there are error correction coefficients in the cointegrated linear equations that ensure that the relationship between variables is stationary $I(0)$, as is the assumption with cointegration (Hill et al., 2018). There are two classes of VECMs: the transitory and the long-term model. We mainly focus on the former in this thesis. In its general form, the transitory VECM can be written as follows:

$$\Delta y_t = \Gamma_1 \Delta y_{t-1} + ... + \Gamma_{p-1} \Delta y_{t-p+1} + \Pi y_{t-1} + \mu + \Phi D_t + \epsilon_t \tag{3.2}$$

In the equation above, we measure the first difference of the ($k$ $x$ $1$)-vector of variables $y_t$, i.e., $\Delta y_t$. $\Gamma_{p-1}$ is a coefficient matrix of lags of the observations in differences, up to the *p-1* lag length, and thus can be thought of as the equivalent of the $A$ matrices in the VAR model.[6] $\Phi D_t$ represents the deterministic trends in the variables. Finally, $\Pi$ is a matrix of the cointegration relationships between the variables, i.e., the error correction.

The matrix of the cointegration relationships can also be defined as $\Pi = \alpha\beta'$, where the dimensions of the matrices depend on the number of cointegrations. Here, the matrix $\beta$

[6]We go up to a lag length of *p-1* because, as stated earlier, the VECM is computed on $p - 1$ lags.

contains the cointegration relationships between the input variables in the model, where the number of columns represents the rank ($r$) of the $k$ number of time series. Meanwhile, $\alpha$ is a matrix that shows the speed of convergence towards the cointegration equilibrium(s). Through restrictions to the $\beta$ and $\alpha$ vectors, one can test hypotheses of a cointegration relationship and the rate of convergence, respectively (Pfaff, 2008). Introducing this notation and making the lagged observations more concise, we can rewrite the equation above as:

$$\Delta y_t = \mu + \alpha\beta' y_{t-1} + \Sigma_{i=1}^{p-1} \Gamma \Delta y_{t-i} + \Phi D_t + \epsilon_t \tag{3.3}$$

This alternative VECM representation wraps up the theoretical concepts section. We will, however, return to these introduced concepts throughout the thesis.

# 4 Methodology

Having familiarised the reader with previous research in the field and key theoretical concepts, we now move on to detailing the specifics of the methodology we will use in order to answer our research questions. This section will first outline the steps necessary to calculate the fundamental price and the theoretical reasoning behind the procedure. This will be the primary focus of this section. Then, when this method has been presented, we will detail how we apply a VECM to analyse the interaction between fundamental and actual price. Overall, our procedure is heavily influenced by the methods of Bergman and Sørensen (2021), as mentioned in the introduction. However, we do deviate on some specific choices and assumptions.

## 4.1 Modeling a Fundamental House Price

In order to determine the fundamental price of the housing market, we begin by interpreting the house price. We start with a simplified example. Consider an agent with rational expectations that plans to purchase a unit of housing in one period, and sell it in the following period. This rational agent should be be willing to pay according to the expected utility of owning and using the unit for this one period. Following Campbell and Shiller, we can deconstruct this utility into three separate factors (1988a).

The first factor is the utility received if the house price is expected to increase after time of purchase, which would provide a utility gain when the unit is sold. The second factor is the cost of owning and using the unit of housing for the period. We follow Bergman and Sørensen in defining this user cost as $\gamma$. The third and final factor, is the utility the agent will get from consuming the housing service. This utility will hereafter be defined as the *imputed rent of housing*. A high utility gained from consumption of housing services will, ceteris paribus, increase the willingness to pay for the unit. Mathematically we can therefore define the agent's willingness to pay for a unit of housing at time $t$ before selling in the future period *t+1* as:

$$P_t = \frac{R_t^H + E_t[P_{t+1}]}{1 + \gamma_t}, \tag{4.1}$$

where the price ($P_t$) is a function of the imputed rent ($R_t^H$) and the expected house price in the next period ($E_t[P_{t+1}]$). Both are discounted with the period's user cost of housing ($\gamma_t$).

Having defined the factors that influence an agent's willingness to pay for housing, we generalize and explain how imputed rents and user costs are calculated in our model. As it serves as the representation of utility, the imputed rent can be thought of as the marginal rate of substitution between housing services and other services and goods and is by Bergman and Sørensen assumed to be represented as shown below in equation (4.2).

$$R_t^H = [i_t(1-\tau_t^i) - \frac{E_t[CPI_{t+1}] - CPI_t}{CPI_t} + \tau + \delta + \eta]P_t - [E_t[P_{t+1}] - P_t] \tag{4.2}$$

Here, $i_t$ is the nominal interest rate, $\tau_t^i$ is the capital income tax rate, $CPI_t$ represents the consumer price index (CPI), while $E[CPI_{t+1}]$ is the expected CPI level for time $t+1$, held at time $t$. Meanwhile, $\tau$ is the effective property tax rate, and $\delta$ is the depreciation rate of the housing stock. Finally, $\eta$ is a premium of owning a unit of housing, due to risk and credit constraints. We note that these last three factors are assumed to be constants and do not vary with time. Here we follow both Bergman and Sørensen (2021) and Hott and Monnin (2008). To ease interpretability of the equation, we remark that $i_t(1-\tau_t^i)$ is the nominal after tax interest rate in period $t$, and $\frac{E_t[CPI_{t+1}]-CPI_t}{CPI_t}$ is the expected percentage change in inflation from period $t$ to period $t+1$. Additionally, $E_t[P_{t+1}] - P_t$ is the expected change in the house price.

Next, we return to the above-mentioned user cost, $\gamma_t$. This represents the relevant costs related to owning and using a unit of housing for a period. Following Bergman and Sørensen, $\gamma_t$ can be neatly defined from equation (4.2) above as the following in equation (4.3).

$$\gamma_t \equiv i_t(1-\tau_t^i) - \frac{E_t[CPI_{t+1}] - CPI_t}{CPI_t} + \tau + \delta + \eta, \tag{4.3}$$

We see that this definition of $\gamma_t$ is can be broken down into two distinct portions. The after tax real interest rate, ( $i_t(1-\tau_t^i) - \frac{E_t[CPI_{t+1}]-CPI_t}{CPI_t}$), where the nominal after tax interest rate is subtracted for expected inflation, and the constants of property tax, depreciation rate, and premium costs.

### 4.1.1 House Price and Forward Iteration

The theory of fundamental valuation assumes that agents are rational and forward-looking. At the start of this section, we introduced an example of an agent owning a unit of housing for one period before selling in the next. This example can, however, be expanded to last several periods, ultimately as far as infinity. This is what Campbell and Shiller (1988a) identified, which has since become the foundation of fundamental pricing. As we have now determined the relationship between house prices, imputed rents and user costs, we consider a new situation where the agent buys the unit of housing in period $t$ but never sells.

In this case, we can refer back to the utility equation (4.2), and see that this implies that the numerator becomes solely the imputed rent ($R_t^H$). Thus, when time goes towards infinity, the unit's value becomes a function of the expected imputed rents and user costs of the period. Formally, this is the concept of forward iteration.[7] Following this principle, we follow Bergman and Sørensen (2021) and define the house price as equation (4.4) below.

$$P_t = E_t[\Sigma_{i=0}^{\infty} \frac{R_{t+i}^H}{\Pi_{j=0}^{i}(1+\gamma_{t+j})}] \tag{4.4}$$

Here, the equation specifically allows for different period-by-period values of the discount factor, user costs. They are based in expectations held at time $t$.

## 4.2 Introducing the Supply and Demand Model

We have now determined how future imputed rents directly influence the house price. However, following our definition of imputed rents as the marginal substitution rate between housing services and other goods and services, future imputed rents thus represent the future willingness to pay for housing services. However, imputed rent values are based in theory, i.e., we do not know concretely what amount of utility it provides, as this is not directly observable. To deal with this issue, we follow Hott and Monnin (2008), who consider a model where imputed rents are used to adjust the supply and demand of housing services with regards to the long-term equilibrium. In other words, $R_t^H$ is used as an adjustment factor. Specifically, we assume the demand for housing services is positively

[7] For further information regarding forward iteration applied to a fundamental pricing of an asset, we refer the reader to Blanchard and Fischer (1989), chapter 5.

influenced by aggregate real disposable income and negatively influenced by imputed rents. Thus, we have the following expression of long-term demand for housing services:

$$D_t = B * Y_t^{\epsilon_Y} * R_t^{H - \epsilon_R}, \tag{4.5}$$

where demand $D$ is determined by a constant $B$, the real aggregate disposable income $Y$, and the imputed rent. The impact of $Y$ on level of demand is influenced by the long-run income elasticity of housing demand, $\epsilon_Y$, while the impact of imputed rent is determined by the price elasticity of housing demand, $\epsilon_R$.

Regarding the supply of housing services, the aggregate housing stock, $H$, is assumed to be equal to the long-term supply of housing services. Since we are considering the situation of a long-term equilibrium, we can make use of the fact that demand equals supply. Therefore, we can substitute the demand for housing services with the aggregate level of housing stock, as seen in equation (4.6).

$$H_t = B * Y_t^{\epsilon_Y} * R_t^{H - \epsilon_R} \tag{4.6}$$

From this, we can isolate imputed rent of housing. This gives us the following equation where $R_t^H$ is defined as a function of the constant $B$, the aggregate real disposable income, the real aggregate housing stock, and the elasticities: $\epsilon_Y$ and $\epsilon_R$.[8]

$$R_t^H = B^{1/\epsilon_R} * Y_t^{\frac{\epsilon_Y}{\epsilon_R}} * H_t^{\frac{-1}{\epsilon_R}} \tag{4.7}$$

The assumption that imputed rents can be used as an adjustment factor in equation (4.6) is an important foundation for the following sections of our methodology. Furthermore, when equation (4.7) is combined with an estimate of fundamental price based on forward iteration as in equation (4.4), essential principles are in place to determine the fundamental house price later on in this section.

[8]We note that we follow Bergman and Sørensen, and assume $B$ to be an unimportant constant as a result of later normalization of the data. As such, $B$ is assumed to be equal to 1

### 4.2.1 Addressing Issues of Stationarity

From section 3.1 on time series analysis, we know that stationarity is necessary to make accurate estimations on time series data. To deal with this issue, we take the logarithm and the first difference in logarithms of the variables, yielding variables that are *I(0)*. We introduce new notation and define the logarithm of housing as $r_t^H$, where $r_t^H = ln(r_t^H)$. We do the same for the other variables where deemed necessary. We also take the first differences, e.g., defining $\Delta r_t^h$ as $\Delta r_t^H = r_t^H - r_{t-1}^H$. This also applies to our other variables. We will use these notations in the following sections.

### 4.2.2 Estimating a Price-to-Imputed-Rent Ratio

This subsection aims to use what we have already defined and calculate an estimate of the price-to-imputed-rent ratio. Such a ratio is beneficial for our analysis as it captures changes to both the house price and the imputed rent level. Thus, it is an efficient ratio for our specific purpose of estimating a fundamental house price. We define the price-to-imputed-rent ratio as follows:

$$S_t \equiv \frac{P_t}{R_t^H} \tag{4.8}$$

Next, we introduce a new notation and let the expectation of the house price in the next period be defined as $P_{t+1}^e$ so that we have: $E_t[P_{t+1}] = P_{t+1}^e$. This notation is also introduced for other variables. This lets us rewrite equation (4.1) from earlier as follows, where we also multiply both sides with $(1 + \gamma_t)$.

$$(1 + \gamma_t)P_t = R_t^H + P_{t+1}^e \tag{4.9}$$

Next, we can introduce the price-to-imputed-rent ratio into the equation by dividing by the imputed rent $(R_t^H)$. Moreover, we substitute $P_{t+1}^e$ with $S_{t+1}^e * R_{t+1}^{He}$, as we can do from equation (4.8) . This results in the following expression:

$$(1 + \gamma_t)S_t = 1 + S_{t+1}^e * \frac{R_{t+1}^{He}}{R_t^H} \tag{4.10}$$

From here, it is possible by the use of logarithms and algebra to show that the logarithm of the price-to-imputed-rent ratio can be written as:

$$s_t = ln(1 + exp(s^e_{t+1} + \Delta r^{He}_{t+1})) - ln(1 + \gamma_t) \tag{4.11}$$

In appendix section (A2), we show how the estimate of $s_t$ in the equation above can be transformed into the the notation we see below in equation (4.12).

$$s_t = c + \Sigma^{\infty}_{j=1} \Psi^j E_t[\Delta r^H_{t+j} - \gamma_{t+j}] - \gamma_t \tag{4.12}$$

Here, there are two notable changes from the prior equations. Firstly, by applying the principle of forward iteration, we represent the logarithm of the price-to-imputed-rent-ratio as an infinite series of differenced imputed rents and user costs. Next, we introduce an adjustment parameter $\Psi$. Here, we follow Bergman and Sørensen (2021). Of note is that they decide to model it endogenously in their model. In our case however, we find it practical to set its value exogenously, as they state has often been the case previously (Bergman and Sørensen, 2021).

With the definition of $s_t$ in (4.12), we return to our supply and demand model. We adjusted the imputed rent so that the supply was equal to the demand of housing services in equation (4.7). We now transform the equation and calculate the differenced log of the imputed rent of housing. This estimate consists of the first differences of the aggregated real disposable income and the aggregate disposable incomes and is as follows:

$$\Delta r^H_{t+j} = (\frac{\epsilon_Y}{\epsilon_R})\Delta y_{t+j} + (\frac{-1}{\epsilon_R})\Delta h_{t+j} \tag{4.13}$$

The last step in estimating our final price-to-imputed-rent ratio is done by substituting (4.13) into (4.12). The result is an estimate of the price-to-imputed-rent ratio that encapsulates forward iteration and adjustments of imputed rents. Importantly, it will serve as a point of departure for the estimation of fundamental price. Below in equation (4.14), we show this new definition of $s_t$. [9]

$$s_t = c + \Sigma_{j=1}^{\infty} \Psi E_t[(\frac{\epsilon_Y}{\epsilon_R})\Delta y_{t+j} + (\frac{-1}{\epsilon_R})\Delta h_{t+j} - \gamma_{t+j}] - \gamma_t \tag{4.14}$$

### 4.2.3 Comments on the Adjustment Parameter $\Psi$

In order to estimate $s_t$, we need an estimate of $\Psi$, where the parameter is defined to be $[0 \leq \Psi \leq 1]$. Through several calculations, Bergman and Sørensen estimate their $\Psi$ parameter as an endogenous parameter in their model. However, as previously mentioned it has been common to estimate this parameter exogenously. As such, we will follow this principle. We base our choice in the following logic.

In the section above, we introduced the price-to-imputed-rent ratio as $S_t = P_t/R_t$, which is also equivalent to: $s_t = p_t - r_t^H$. In a long-term equilibrium then, we will expect the equation to hold. An approximation of this long-term value is the mean value of the variables. Then we get the following relationship:

$$\overline{s} = \overline{p} - \overline{r^H} \tag{4.15}$$

Bergman and Sørensen also point out that they calibrate the imputed rent so that:

$$\overline{r^H} = \overline{p} + log(\overline{\gamma}) \tag{4.16}$$

Substituting this definition into equation (4.15) above, we find that one can assume that $\overline{s} = log(\overline{\gamma})$. Following this reasoning, we decide to calibrate $\Psi$ such that the the mean price-to-imputed-rent ratio equals the logarithm of the mean value of user cost. However, as the logarithm of low percentages would return a considerably large value, and

[9] We note that this equivalent to Bergman and Sørensen (2021) and their equation 6.

considering some ambiguity from Bergman and Sørensen, we make the assumption that

$$\Psi = log(1 + \overline{\gamma}) \tag{4.17}$$

Regarding the robustness of our estimate of $\Psi$, the choice to estimate $\Psi$ exogenously might provide limitations to our $\Psi$ estimate. This is something that will be addressed in both the analysis and the discussion section.

## 4.3 Applying a VAR Model Framework

Now that we have calculated an estimate of the price-to-imputed-rent ratio, this section aims to explain the necessary steps of applying the VAR model. The estimated coefficients can be used to calculate the fundamental house price based on concepts introduced in the theory section (3.2).

A first step, is to determine the input variables, i.e., what variables we are determining the coefficient of. The VAR model must, at a minimum, include the variables introduced earlier that determines the fundamental house price. Therefore, we will include time series of the differenced logarithm of aggregated real housing stock ($\Delta h_t$), the differenced logarithm real disposable income ($\Delta y_t$), and the user costs of housing ($\gamma_t$). Bergman and Sørensen also include the differenced logarithm of actual house prices ($\Delta p_t^a$) in their model. This is further backed by Campbell and Ammer (1993) and Engsted et al. (2012) who underpin the importance of including the actual price when estimating how fundamental factors influences stock price. Therefore, we include actual price in our VAR model. Lastly, Bergman and Sørensen consider the differenced logarithm of real housing rent costs ($\Delta r_t$).[10] This rent cost variable is not strictly needed from theory or as a direct input in calculating fundamental price. However, Bergman and Sørensen include it as a useful proxy to a host of other variables that might influence the level of house prices, and the estimated coefficients of the other variables.

[10]This is not to be confused with the aforementioned imputed rent of housing ($r_t^H$), which is the unobservable utility of using some unit of housing.

Thus, we follow Bergman and Sørensen (2021) in their choice of variables, and rely on a five variable VAR model. Following this, we introduce $b_t$ below in equation (4.18). This is a time series vector consisting of the VAR variables, and is our model's definition of $y_t$ from the VAR model defined in section (3.2) on VAR models.

$$b_t \equiv \begin{bmatrix} \Delta p_t^a \\ \Delta r_t \\ \gamma_t \\ \Delta y_t \\ \Delta h_t \end{bmatrix} \tag{4.18}$$

### 4.3.1 Implementation of VAR Model

When applying the VAR model, the intent is to estimate the explanatory effect the previous observations might have on the present levels of the variables. In equation (4.19) below, our $\Phi$ is the equivalent to the $A$ in equation (3.1) in the theory section. In our case, each $\Phi$ represents a $(5\ x\ 5)$-matrix containing the coefficients of the variables in each lag, up to the lag length $p$. The error term, $\epsilon_t$, is a $(5\ x\ 1)$-column vector. In matrix notation, our VAR(p) model can be written as follows:

$$b_t = c + \Phi_1 b_{t-1} + \Phi_2 b_{t-2} + ... + \Phi_p b_{t-p} + \epsilon_t \tag{4.19}$$

In section (A2.3) in the Appendix, we also show how equation (4.19) can be written more concisely as equation (4.20).

$$\Phi(L) b_t = \epsilon_t \tag{4.20}$$

### 4.3.2 Modeling Expectations from the VAR Output

After defining the five variable VAR model(p), the next step in estimating the fundamental house price is computing an interpretation of expectations. This is needed in order to specify the price-to-imputed-rent ratio, $s_t$, as shown in equation (4.14). This subsection will outline the necessary steps to arrive at such an estimate, using the output of the VAR(p).

We begin by defining the vector $z_t$, as a collection of the variable vectors $b_t$, $b_{t-1}$, up to $b_{t-p-1}$. $z_t$ is defined below. This notation is convenient, as $z_t$ can be estimated purely by $z_{t-1}$, as it contains the correct number of lagged observations, as defined by lag length $p$.

$$z_t \equiv \begin{bmatrix} b_t \\ b_{t-1} \\ . \\ . \\ b_{t-p+1} \end{bmatrix} \tag{4.21}$$

From the properties of $z_t$, we can thus rewrite our VAR(p) model as a VAR(1) model, also called the companion-form (Kotze, 2022).[11] From this definition we can rewrite (4.21) as:

$$z_t = Az_{t-1} + \xi_t \tag{4.22}$$

We see that $z_t$ is defined by $z_{t-1}$, i.e., the $p$ number of lagged observations, multiplied with $A$, a companion matrix containing the coefficient estimates $\Phi$ of the lag lengths. The companion matrix can be shown to be of dimension ($kp$ $x$ $kp$), and as number of variables $k$ equals five, we have that the companion matrix is of dimension ($5p$ $x$ $5p$). Meanwhile, $\xi$ is a ($p$ $x$ $1$)-matrix containing only the error terms of the equation for the current period. This is because we do not need the error term of previous periods, as we have the actually observed values. We rewrite (4.22) to the following format in equation (4.23) to better visualise the definition of $z_t$.

[11] We note that according to Kotzé, Bergman and Sørensen, and others, we can always rewrite a VAR(p) model as a VAR(1) model

$$z_t = \begin{bmatrix} \Phi_1 & \Phi_2 & . & \Phi_{p-1} & \Phi_p \\ I_5 & 0 & . & 0 & 0 \\ 0 & I_5 & . & 0 & 0 \\ . & . & . & . & . \\ 0 & 0 & . & I_5 & 0 \end{bmatrix} \begin{bmatrix} b_{t-1} \\ b_{t-2} \\ . \\ . \\ b_{t-p} \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ 0 \\ . \\ . \\ 0 \end{bmatrix} \tag{4.23}$$

At this point, we follow Bergman and Sørensen in assuming the expected value of a variable in the model from (4.22) can, in the general case, be rewritten as:

$$E_t[z_{t+j}] = A^j z_t \tag{4.24}$$

Specifically, it can be shown to be the case for $j = 1, 2, ...n$ periods ahead. To help interpretation we provide step-by-step calculations for the instances of $j = 1$ and $j = 2$. In the case of $j = 1$, we get the equation (4.25).[12] Meanwhile, equation (4.26) represents the case of $j = 2$.

$$E_t[z_{t+1}] = Az_t \tag{4.25}$$

$$E_t[z_{t+2}] = E_t[E_{t+1}[z_{t+2}]] = E_t[Az_{t+1}] = AE_t[z_{t+1}] = A^2 z_t \tag{4.26}$$

It follows from this mathematical definition, that we can model expectations of the future as a function of the estimated coefficients tied to the lags of the variables, based on today's ($z_t$) values. However, we note that in order to compute the companion matrix, and thus the VAR model, must be stable. In effect, this implies that the eigenvalues of the VAR model should be inside the unit circle, i.e, less than one in absolute terms. A stable VAR model also implies full rank, i.e., the number of cointegrations is such that $r = k$. We will test for these assumptions in the analysis.

[12]The error term is not included in the above equation. This is because the expected value will only be measurable when looking backwards.

## 4.4 Final Steps Towards the Fundamental Price

### 4.4.1 Quantifying the Price-To-Imputed-Rent Ratio

We have now shown how to define the price-to-imputed-rent ratio as required by the supply and demand model of Bergman and Sørensen, and how we will apply a VAR(p) framework in order to model expectations. In this subsection we will tie these pieces together and arrive at an estimate of fundamental house price, which of course is a prerequisite in order to answer our research questions. This section is closely aligned with the procedure laid out by Bergman and Sørensen.

We start this section by introducing a pair of column vectors, $g_1$ and $g_2$. The dimensions of the vectors are (*1 x 5p*), with column length decided by lag length $p$. The definition is shown below in equation (4.27) and (4.28).

$$g_1 \equiv \begin{bmatrix} 0 & 0 & -1 & \frac{\epsilon_Y}{\epsilon_R} & \frac{-1}{\epsilon_R} & 0 & . & 0 \end{bmatrix} \tag{4.27}$$

$$g_2 \equiv \begin{bmatrix} 0 & 0 & -1 & 0 & 0 & 0 & . & 0 \end{bmatrix} \tag{4.28}$$

The position of the values in the vectors corresponds to the position of the variables in our VAR(p) model. For instance, $g_2 * z_t = (-1) * \gamma_t$. This is because $\gamma_t$ is located in the third row of equation (4.18). Using these characteristics of $g_1$ and $g_2$, we can reintroduce the definition of $s_t$ in equation (4.14) and redefine the price-to-imputed-rent ratio as:

$$s_t = c + \Sigma_{j=1}^{\infty} \Psi^j g_1 E_t[z_{t+j}] + g_2 z_t \tag{4.29}$$

Next, we rewrite the expression by implementing the new definition of future expectations as a function of the companion matrix $A$, as shown in (4.24). Furthermore, we remove the constant $c$, which according to Bergman and Sørensen is not important for present purposes. From this we then get:

$$s_t = \Sigma_{j=1}^{\infty} \Psi^j g_1 A^j z_t + g_2 z_t \tag{4.30}$$

Which is equivalent to:

$$s_t = [(\Sigma_{j=1}^{\infty} \Psi^j g_1 A^j) + g_2] z_t \tag{4.31}$$

Following Bergman and Sørensen (2021), we rewrite this equation to arrive at the final representation of the price-to-imputed-price ratio, shown below in equation (4.32).

$$s_t = [g_2 + \Psi g_1 A (I - \Psi A)^{-1}] z_t \tag{4.32}$$

With this definition, we have all necessary parameters to calculate $s_t$ in quantifiable sizes.

### 4.4.2 Computing the Fundamental Price

With the price-to-imputed-rent ratio now defined in a way that allows us to calculate its period-by-period value, we can finally turn to the main issue at hand, estimating the fundamental house price. In order to arrive at a mathematical construct of the price, we make use of the definition $s_t \equiv p_t - r^H$, which can be seen from equation 4.8.[13] Changing the expression, we see that we can define price as $p_t \equiv s_t + r^H$. By substituting the log of imputed rent, $r^H$, with its definition in equation (4.7), we arrive at the final, quantifiable expression of fundamental price in equation (4.33).[14]

$$\hat{p}_t^f = s_t + ((\frac{\epsilon_Y}{\epsilon_R}) y_t - (\frac{1}{\epsilon_R}) h_t) \tag{4.33}$$

We will employ this equation in order to estimate a fundamental valuation of the Norwegian housing market, and to evaluate the estimate against the observed levels of real house prices.

[13]Because $S_t = P_t / R_t^H, s_t = ln(P_t / R_t^H) = p_t - r_t^H$

[14]Given the previous assumption of $B = 1$

## 4.5 Modeling the Relationship between Real and Fundamental Price

### 4.5.1 Estimating the VECM and Assessing Cointegration

Having presented the method necessary to estimate the fundamental house price, we turn to our second research question, i.e., the interaction between the actual and fundamental house prices. We will show how we implement a VECM, introduced in section (3.3), in order to determine if there is such a relationship, through a test of cointegration.

To do this, we introduce a vector of both the fundamental and actual price, defined as the following: $q_t \equiv [p_t^a, p_t^f]'$. As stated in the theoretical concepts in section (3.3), it is always possible to rewrite a VAR model as a VECM, provided it has a lag length $p \geq 2$ (Pfaff, 2008). With this as a basis, we begin by following Bergman and Sørensen and assume $q_t$ can be generated through a VECM, where both time series are $I(1)$. This general notation is shown in equation (4.34).

$$\Delta q_t = \Gamma_1 \Delta q_{t-1} + ... + \Gamma_{p-1} \Delta q_{t-p+1} + \Pi q_{t-1} + \mu + \Phi D_t + \epsilon_t \tag{4.34}$$

Simplifying the notation above, we note that we follow Bergman and Sørensen in removing all deterministic trends, i.e., $\Phi D_t$. Next, we can rewrite the $\Pi$ representation from this theoretical approach as $\alpha\beta'$. Summarizing the lags, we get the following, more specific VECM for our purposes:

$$\Delta q_t = \mu - \alpha\beta' q_{t-1} + \Sigma_{i=1}^{p-1} \Gamma \Delta q_{t-i} + \epsilon_t \tag{4.35}$$

Here, $\Delta q_t$ represents the first difference of the *(2 x 1)* $q_t$ vector of the time series. We also have the *(2 x 1)*-vector of constants, $\mu$, as well the error term, $\epsilon_t$, of the same dimension. Meanwhile, $\Gamma$ will contain the estimated coefficients of the VECM. The loading matrices $\alpha$ and $\beta$ have dimensions conditional on the rank $r$, and conitaing the cointegration relationship(s) and thus the error correction. When applying this model in the analysis, the first issue at hand will be assessing the interaction between the fundamental and actual price. We will test this by analysing the number of cointegrations, i.e., long-term relationships between variables as explained in section (3.2.3). In order do to this, we will utilise the Johansen Trace Test for cointegration (Johansen, 1991). Applying this test on the estimated VECM, we will test for the validity of different values of rank $r$. A cointegration rank of one between $p_t^f$ and $p_t^a$ would imply a cointegrated relationship. We will come back to this in the analysis section.

### 4.5.2 Restrictions on the $\alpha\beta$ Loading Matrices

We now turn to the next aspect of our second research question. We are not only interested in evaluating if there is a cointegration relationship between actual and fundamental price, but also if the relationship is consistent with the theory of fundamental valuation.

The first assumption that we will test for is the assumption of a stationary gap between actual and fundamental prices. Effectively, this implies that $p_t^a - p_t^f \sim I(0)$, as defined by Bergman and Sørensen (2021). Thus, the gap has to be mean reverting, as if it were not, it would have a trend component and therefore not be stationary. We test for this specific assumption by use of the VECM defined in (4.35). We will restrict the cointegration matrix $\beta$ such that $\beta' = [1, -1]$. This restriction implies that a gap between the two time series will be adjusted by the error correction term.

The second assumption we introduced in our introduction, was the assumption that only actual house price would react to a gap between the prices. Tying into the assumption above, this implies that only the actual price is actively mean reverting in the error correction term. We can test the validity of this assumption through further restrictions to the VECM, this time to the adjustment coefficient matrix $\alpha$. From our definition of $q_t$, we have that the first row in the vector is actual price ($p_t^a$), and the second is the fundamental price ($p_t^f$). Therefore, we will test for restrictions to the adjustment coefficient matrix such that $\alpha_1 < 0$ and $\alpha_2 = 0$. These restrictions thus assures that the actual price should react negatively to a gap between the two prices, while an adjustment coefficient value of zero for the fundamental price should ensure that the price estimate is modeled independently of any gap between the series.

We will in the analysis test whether these assumptions, i.e., restrictions, return a VECM that can plausibly describe a relationship between the two time series, given that there is a cointegration relationship between them.

# 5 Data

So far, we have introduced the aim of this thesis, discussed previous literature on the subject, introduced relevant theoretical concepts and laid out our chosen methodology. The following section introduces the data on which we apply our methodology. Since this thesis aims to create a representative analysis of the Norwegian housing market, this has been the focus when gathering necessary data.

Our primary data source is Statistics Norway and their statistics on the Norwegian national accounts. Most of the data is taken from Statistics Norway's database *KVARTS*, a large macroeconometric model used to analyze quarterly business cycle forecasts and policy analyses in the short and medium term (SSB, 2022c). KVARTS data is available upon request. The main data we need to estimate a fundamental price, are the five variables specified in our methodology as inputs for the VAR model. The variables were presented in equation (4.18), where we defined the vector $b_t$ as $b_t' \equiv \begin{bmatrix} \Delta p_t^a & \Delta r_t & \gamma_t & \Delta y_t & \Delta h_t \end{bmatrix}$.

Therefore we requested data on the actual housing price, rent price of housing, disposable income, and housing stock. We also requested data on the real interest rate after tax, which we need to calculate the user cost, and the Consumer Price Index (CPI). The variables are all provided on a quarterly basis and are not seasonally adjusted. They cover the period from Q1 1982 to Q2 2022. Where appropriate, we have transformed the data into real terms with 2019 as the base year (i.e., with a value of 100). In addition, because we are working with macroeconomic data, we calculate the logarithm and the first difference of the logarithm for some of the variables. [15] We will now describe each variable and the steps we took to transform the data into a uniform format suitable for our analysis.

[15]A detailed overview of our data set and its descriptive statistics can be found in the appendix in tables (A3.1) and (A3.2), respectively.

## 5.1 Inflation and Population Variables

As mentioned, our data set includes variables of various types, including nominal and real values as well as aggregated and indexed data. In order to make our analysis more straightforward, we need to transform the data into a more uniform format. We will first introduce the necessary components in such a transformation before moving on to the main variables in our analysis.

**Consumer Price Index**

The consumer price index (CPI) is a measure that describes the price development of goods and services that private households purchase and is thus a measure of inflation (SSB, 2022a). Since some of our variables are in nominal terms, the CPI deflates the nominal values to real ones. Here we specifically follow the principle that nominal value / CPI = real value. CPI data comes from the KVARTS database and has the average of the four quarters of 2019 as the basis, equaling 1.

**Population Data**

Data on Norwegian population levels are gathered from Statistics Norway's table 06913, containing the yearly population in Norway from 1951 to 2022. We transform the data to a quarterly basis under the assumption that the population growth is evenly spread out throughout the quarters.

## 5.2 Introducing the Main Variables

**Actual House Price**

Statistics Norway calculates the quarterly house price as a nominal index in their KVARTS database with the average of the four quarters of 2019 as the basis, equal to 100. The index represents the development of the aggregated Norwegian housing market and contains the change in the price of apartments and one- and two-family dwellings per quarter from Q1 1982 to Q2 2022. For our analysis, we require variables in real terms and therefore deflate the nominal index by the CPI for each observation. In addition, we choose to seasonally adjust the data of actual house prices. When comparing adjusted to unadjusted data, we see clear seasonal trends - potentially reducing the preciseness of our econometric analysis.

Moreover, this thesis aims to provide an overview of the Norwegian housing market and its overall long-term trends rather than monitoring the specific period-by-period situation. We also take the logarithm of each observation and calculate the period-by-period change in the logarithm.

**Figure 5.1:** Real Actual House Price: Logarithm and First Difference

**(a)** Logarithm of Actual Real House Price. **(b)** First Difference of Actual Real House Price.

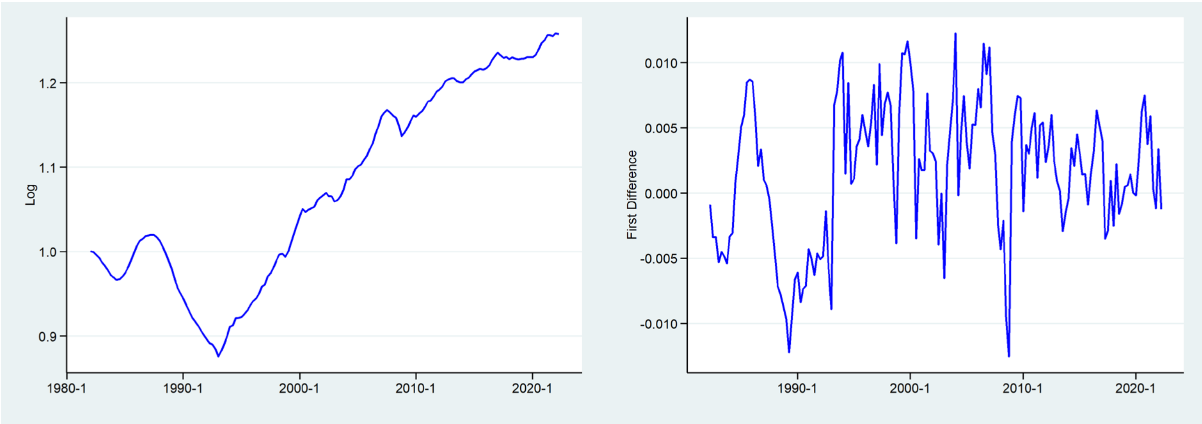


Figure 5.1 displays the logarithm of the real actual house price and the period-by-period change of the logarithm of actual house price for Q1 1982 to Q2 2022. The logarithm of actual house price is in the first period normalized to one. So does Bergman and Sørensen in their figure 1 2021.

We plot the logarithm of actual house price in Norway in figure (5.1a) above. We can see that prices decreased substantially at the beginning of the sample period, hitting bottom around 1993. However, from this point onward, there has been a steady increase in real house prices over the sample. Regarding the period-by-period change in the logarithm of the actual house price in figure (5.1b), we see more marked trends in the first part of the sample. From then on, the mean of the time series seem relatively consistent across the periods, although one could argue that there has been a decrease of the variance since 2010.

Our data on actual house price differs from Bergman and Sørensen (2021), who only include one- and two-family dwellings in their analysis, thus excluding apartments from their sample. We argue that utilising data on the whole Norwegian housing market could lead to a more comprehensive analysis, as apartments represent a substantial proportion of the housing supply in Norway. Moreover, not including apartments could skew the results when comparing housing supply to other variables.

### Rent Prices of Housing

Our house rent price data also stems from the KVARTS database. The data shows the development in the average paid house rent in the Norwegian market for a unit of housing. Structurally, it is a nominal index of quarterly data, with the average of the four quarters of 2015 as its basis, equal to 100. To compute the index in real terms, we deflate by CPI. After that, we transform the data to 2019 = 100. This is done by dividing each observation by the average value of the 2019 observations before multiplying all observations by 100. We also transform it by taking the observations' logarithm and calculating the first difference in logarithm as with the other variables. Below we have plotted the evolution of the log of the rent variable and the first difference change in the logarithm of rent prices in figure (5.2). In figure (5.2a), we see the same downward trend in the early observations as in the house prices. However, the upturn after this point is markedly more jagged, and after 2016 we see a fall in real rent prices of housing in Norway. In the same period, we see that real house prices continued their upwards trajectory. Commenting on figure (5.2b), we see no trajectory or change in variance over the sample.

**Figure 5.2:** Real House Rent Price: Logarithm and First Difference

**(a)** Logarithm of Rent Price of Housing. **(b)** First Difference of Rent Price of Housing.

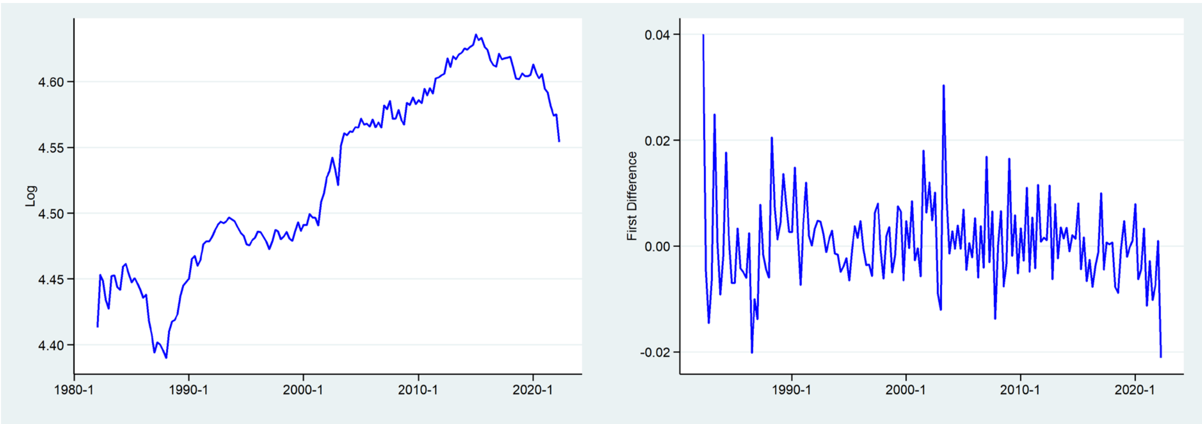


Figure 5.2 shows the logarithm of the rent price of housing in Norway and the period-by-period change in the logarith of the rent price. The data spans Q1 1982 - Q2 2022.

### Stock of Housing

The estimate of the Norwegian housing stock is the aggregated real capital housing stock from Statistics Norway found in the KVARTS database. We have data from Q1 1982 to Q2 2022 showing the real capital allocation in Norway in NOK million. To find the housing stock on a per capital level we divide the aggregated data with the population data, before taking the logarithm and calculate the first difference of the logarithm. We see that the housing stock steadily increases throughout the entire sample period, as shown in figure (5.3a). The data is otherwise stable, and does not seem to have much volatility over the sample period. However, in figure (5.3b) we see that the period-by-period change of the logarithm is marked by clear trends in the growth rates. There is also a period of seemingly higher volatility in growth rates at the beginning of the sample, which seems to dissipate around 1990.

**Figure 5.3:** Real Housing Stock: Logarithm and First Difference

**(a)** Logarithm of Housing Stock. **(b)** First Difference of the Housing Stock.

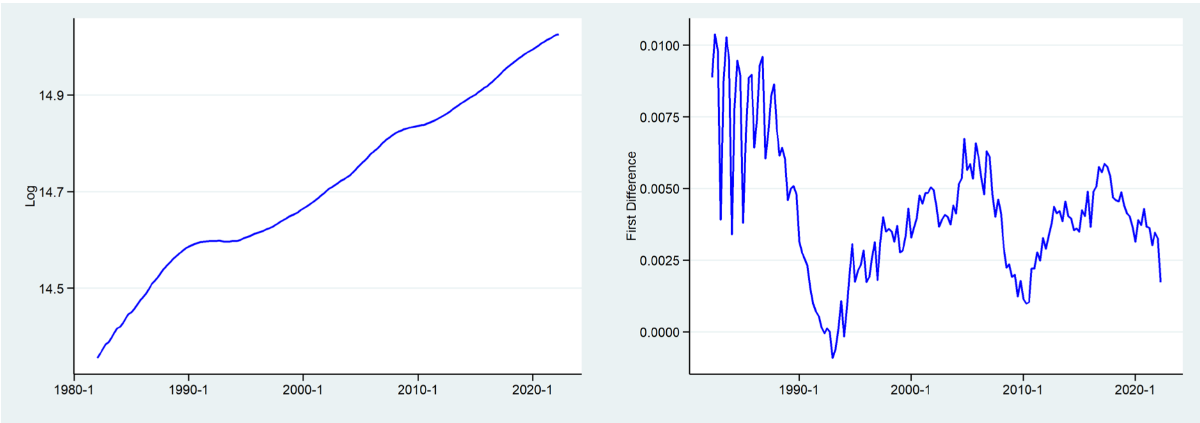


Figure 5.3 shows the logarithm of the housing stock in Norway per capita and the period-by-period change in the logarithm of the housing stock for the sample of Q1 1982 - Q2 2022.

### Disposable Income

We also need data on levels of disposable income over the sample period. Here we turn back to the KVARTS database. The data is the aggregated real disposable income of Norwegian households from the national accounts in NOK million from Q1 1982 to Q2 2022. For this thesis, however, the real disposable income per capita is a better choice. Therefore, we divide the observations on the population data. Moreover, we seasonally adjust the variable. When examining adjusted and un-adjusted data, the latter contained repeating seasonal patterns, perhaps influenced by remuneration like yearly bonus payments. As a result, this data is less stable than seasonally adjusted data. Another argument for seasonally adjusting the variable is that households probably are more likely to plan their purchasing decisions for significant purchases with a basis on an average disposable income over a longer period, rather than the disposable income of a specific quarter. Analysing the logarithm of disposable income shown in figure (5.4a), we can see that it increases steadily over the entire sample. However, the period-by-period change in the logarithm in figure (5.4b) does not seem to indicate any clear trend with a similar mean and standard deviation across the sample.

**Figure 5.4:** Real Disposable Income: Logarithm and First Difference

**(a)** Logarithm of Disposable Income. **(b)** First Difference of Disposable Income.

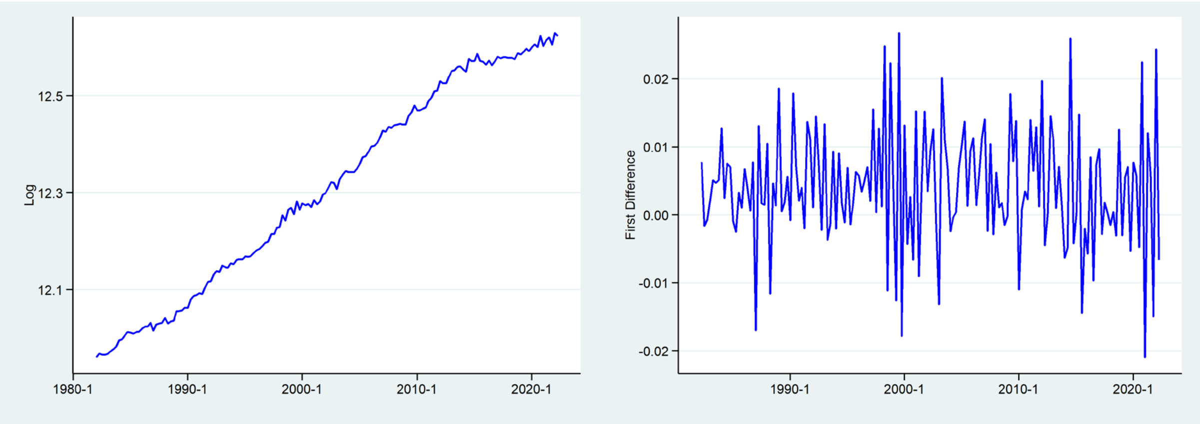


Figure 5.4 shows the logarithm of the disposable income in Norway and the period-by-period change in the logarith of the disposable income.

### User Cost of Housing

A critical variable in determining the fundamental house price is the user cost of housing. This variable is determined as a function of the real interest rate after tax, depreciation of the housing stock, effective property tax, and risk premium, as laid out in equation (4.3) in the methodology section.

The real interest rate after tax variable comes from the KVARTS database and is the average nominal interest rate paid by Norwegian households per quarter, deflated with CPI to real terms. The after-tax component is calculated by subtracting the tax benefit associated with paying interests, which is determined by the capital- and interest-tax rate. Because of the tax benefit, the after-tax interest rate is lower than the pre-tax rate.

To compute the user cost, we also need an estimation of the rate of depreciation of the housing stock, which is a proxy for the quarterly necessary investment costs to maintain a unit of housing for a quarter. The depreciation rate, $\delta$ is calculated based on yearly data from Statistics Norway on housing depreciation and total real capital in housing from 1970 to 2021 from their table 09181, with real values indexed to 2015 = 100. The depreciation rate can also be shown in the following way:

$$Dep.rate_t = \frac{Dep_t}{Cap_{t-1}} \tag{5.1}$$

Here, $Dep.rate_t$ is the depreciation rate of the housing stock at period $t$, $Dep_t$ is the aggregated depreciation of the Norwegian housing stock at year $t$, and $Cap_{t-1}$ equals the aggregated real capital stock in the Norwegian housing stock of the previous period. Calculating the average depreciation rates across our sample period yields an average rate of $2,1155\% \approx 2\%$ on a yearly basis. We assume this depreciation rate to be representative for the entire sample period. Therefore, with a basis in our calculations we assume a constant depreciation rate of the housing of $(1.02)^{1/4} - 1 \approx 0.5\%$ per quarter.

Next, we also need to estimate the effective property tax. Bergman and Sørensen find this to be a challenging component to quantify, as is the situation in our case. We have not been able to find any existing data for the effective property tax for the aggregate of the Norwegian households, and a large variety of factors might plausibly influence it. As such, this paper follows Bergman and Sørensen in assuming that the effective property tax rate is constant across the sample. In this regard, our chosen proxy is the mean of the yearly paid property tax from housing in Norway divided on the total nominal capital allocation of housing in Norway. The data is available yearly from 2007 to 2021 in Statistics Norway's table 06980 and 09181, respectively. This yields an estimated average property tax of 0.143%, approximately corresponding to a quarterly rate of $(1.00143)^{1/4} - 1 \approx 0.035\%$ per quarter from 2007 to 2021. Although it should be noted that the property tax as we know it in Norway today was implemented in 2006, there have been other similar taxes in Norway in earlier periods (Bergsholm, 2022). We assume that these former tax policies are of approximately the same magnitude as the existing tax regime, and that our calculated effective tax rate is constant and representative for the whole sample period.

The last variable needed to calculate the user cost is the risk and capital constraint premium, as seen in equation (4.3). Bergman and Sørensen (2021) assume this to be constant for their sample. We assume it to be zero across our sample. This is equivalent to assuming agents are risk-neutral and are not significantly influenced by credit constraints. Thus, we can compute the user cost described in the methodology section. Below, the user cost of each quarter of the sample is plotted in figure (5.5). As we can see, the user cost peaked in 1993, rising sharply from low levels at the start of the sample. Since the peak, user costs have slowly decreased, and are, as of last observation Q2 2022, equal to $-3\%$ , mainly driven by a negative real interest rate component.

**Figure 5.5:** Real User Cost of Housing.

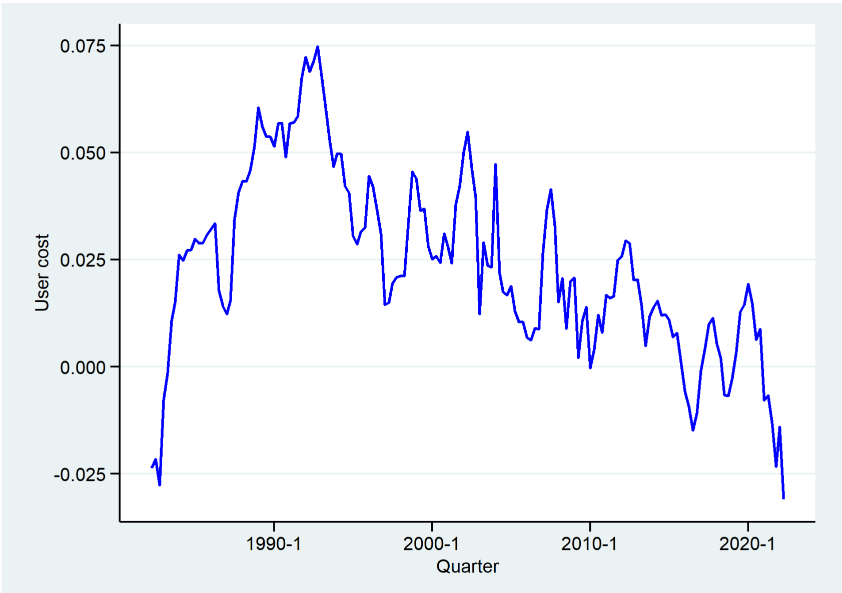


Figure 5.5 shows the real user cost of housing for the Norwegian households on a quarterly basis from Q1 1982 to Q2 2022. The user cost consists of the after tax real interest rate, depreciation rate of housing stock, effective property tax and risk- and capital premium.

# 6 Analysis

In this part of the thesis, we aim to answer our research questions by first estimating the fundamental house price in Norway and testing the robustness of our estimate. Thereafter, we will analyse whether there is a stable relationship between the fundamental house price and the actual price and whether this potential relationship is in accordance with what the theory of fundamental valuation suggests. For the specifics of our computation, we will primarily rely on the R-packages *vars* and *urca* for VAR and VECM, respectively. Both packages are authored by Pfaff (2008)

## 6.1 Addressing a Potential Structural Break

As presented in the methodology section, a prerequisite for analysing time series with an autoregressive model is a stable mean and variance of the variables throughout the sample. Deviation from this can indicate a structural break in the data, i.e., that some variables behave differently in different periods. Concerning a VAR model, this would indicate that the coefficients would be significantly different if estimated on different subsets of the data. This is problematic, as different behaviour of the variables in subsets of the data will likely cause some adjustment of the coefficients. The result is that the model loses accuracy in estimating the actual behaviour of the variables since the model will attempt to create one uniform estimate (Antoch et al., 2019). When dealing with macroeconomic variables, reasons for such structural breaks can be a change in collective preferences, tightening or loosening of the regulatory environment or major geopolitical events, to name a few (Hailegiorgis et al., 2011).

Assessing the plots of our variables from the previous section on data, we see indications that we have a structural break in our data set. Figure (5.1) indicates that the behaviour of actual house prices changed notably in the early 1990s. We also note that when taking the first differences of the observations in (5.1b), there are indications that the early observations exhibit a distinguishable pattern. This seems to dissipate towards the mid-90s. Furthermore, figure (5.3) indicates an apparent change in the variance and the mean of the housing stock differences. Finally, our calculated variable of user costs also indicates a break. Figure (5.5) shows how user cost increased towards the early 1990s, reaching a peak in Q1 1993, before trending steadily downwards for the remainder of the sample period.

Comparing these observations with the housing market at the time, it has been reasonably established that the Norwegian housing market had bubble-like characteristics at that time, bursting in the early 90s (Eitrheim et al., 2004). The period also coincides with a loosening of government regulations on housing and financial markets, as mentioned by Moe et al. (2004). These circumstances add additional weight to the argument of a meaningful structural break in our data. Furthermore, dealing with bubbles in data is particularly important from an econometric point of view, as they are clear breaks that do not explain changes in the fundamental variables (Escobari et al., 2015).

Comparing our data with Bergman and Sørensen, they state that they found a clear structural break in their data set due to the Swedish housing bubble in the early 1990s, followed by a banking crisis lasting until 1994. They find these factors to lead to exceptionally volatile housing prices and user costs for the period. They offer two solutions, i) estimating a model which can distinguish between the housing markets before and after this structural shift, or ii) excluding this period from the sample and estimating the fundamental price starting after this period. They choose to exclude the period from their estimation of the fundamental pricing model.

### 6.1.1 Impact for our Analysis

Although there are drawbacks to voluntarily reducing the size of a data set when estimating a model, the arguments in favour of excluding the early part of our data set from our estimation are well founded. Thus, we follow Bergman and Sørensen and will not use our entire sample. Specifically, we estimate our VAR model on data from Q2 1993, which corresponds to the quarter after the user cost peaked and the house prices bottomed out. We will later on test this assumption and analyse the implications of estimating with the entire data set.

## 6.2 Estimating the VAR Forecasting Model

### 6.2.1 Assessing the Specifications of the VAR(p) Model

To calculate the VAR(p) model, we will use the *vars* package in R, as suggested by Pfaff (2013). This will allow us to estimate the coefficients needed to compute the price-to-imputed rent ratio $s_t$, which is a key component of the fundamental price calculation outlined in the methodology section. As part of this process, we will first need to determine the optimal lag length for our model, as we will be including lagged versions of the variables in our analysis.

For this purpose, we use the function *VARselect()*, which returns four estimates of the optimal lag length for a VAR(p) model based on four different information criteria; Aikake information criterion (AIC), Hannan-Quinn criterion (HQ), the Final Prediction Error criterion (FPE), and Schwarz criterion (SC), also known as the Bayesian Information Criterion (BIC). The test results are shown in table (6.1). We find that the AIC and FPE criteria indicate an optimal lag length of five, while HQ and SC criteria return a length of one. The optimal lag length should have no issues in the residuals. Using the HQ and SC prediction as a lower bound, we use an iterative process starting at one lag, where we test for autocorrelation, heteroscedasticity, and normality of the residuals. If these tests indicate significant issues with the residuals, we add one more lag and re-estimate the VAR(p) model.

**Table 6.1:** Testing For Optimal Lag Selection for a VAR(p) Model

| Test | AIC(n) | HQ(n) | BIC(n) | FPE(n) |
|---|---|---|---|---|
| Lags | 5 | 1 | 1 | 5 |

Table (6.1) displays the four different lag options and their respective proposed lag lengths. We consider the different information criteria: Aikake information criterion (AIC), Hannan-Quinn criterion (HQ), the Final Prediction Error criterion (FPE), and Schwarz criterion (SC). $n$ is the number of lags the tests are evaluated in, which in this case, a maximum lag length of 12.

Using this process, we find that a lag length of five is desirable for our sample period. This is in accordance with the optimal lag length indicated by the AIC and FPE criteria, which is reassuring. The residual tests for this VAR(5) model are plotted in table (6.2). Regarding autocorrelation, we can see that when estimating the VAR model with five lags, the null hypothesis of no autocorrelation of the Breusch-Godfrey test cannot be rejected as we get a p-value of 0.341. There are no issues with heteroscedasticity either, as a p-value of 0.4462 from the ARCH test indicates that our residuals have little issue with changes to the mean and variance across the sample. Lastly, the VAR model seems to fulfil all requirements of multivariate normality of the residuals, with a p-value of 0.4471 of the Jarque-Berra multivariate normality test and exhibiting no clear indications of kurtosis or skewness.

**Table 6.2:** VAR Residual Tests for Autocorrelation, Heteroscedasticity and Normality

| | Autocorrelation | Heteroscedasticity | Normality | | |
|---|---|---|---|---|---|
| Test | LM(5) | ARCH(5) | Jarcue-Bera | Skewness | Kurtosis |
| p-value | 0.341 | 0.4462 | 0.4471 | 0.3103 | 0.5542 |

Table (6.2) displays the results of our Breusch-Godfrey test of autocorrelation, the ARCH-test of heteroscedasticity, the Jarque-Bera test for normality, as well as multivariate tests for skewness and kurtosis.

### 6.2.2 Granger Causality Test of the VAR Variables

Going further, we decide to conduct a Granger Causality test on our VAR(5) model. As mentioned in the theoretical framework, this should not be mistaken for actual causality but as predictive causality. Therefore we instead refer to variables as "Granger-causing" other variables. Applying the function *causality()* from the *vars*-package, we run the tests and list the results in table (6.3). Here the null hypothesis is that a variable does not Granger-cause the other variables. The table shows that the variables "Real price" and "Real rent price" are evaluated as significant in estimating the other variables at a 5% confidence level. This is somewhat outside what we would expect as, per our methodology, we assume that it will be the fundamental factors $\gamma_t$, $\Delta y_t$, and $\Delta h_t$, which would cause changes to real price ($p_t^a$) and not the opposite. [16]

[16]Regarding one variable Granger-causing the other variables, we note that we cannot separate which specific variables are Granger-caused by the variable.

**Table 6.3:** Granger Causality Tests for our VAR(5) Variables

| | Variable | p-value |
|---|---|---|
| $\Delta p_t^a$ | Real price | 0.03904** |
| $\Delta r_t$ | Real rent price | 0.01738** |
| $\gamma_t$ | User cost | 0.2614 |
| $\Delta h_t$ | Housing stock | 0.7539 |
| $\Delta y_t$ | Disposable income | 0.1797 |

* p<0.10, ** p<0.05, *** p<0.01

Table (6.3) displays the Granger causality tests with the null hypothesis that a given variable does not Granger-cause the other variables.

### 6.2.3 Considering the Stability of the VAR(5) Model

The theoretical procedure by which we estimate our VAR(5) model is based on the assumption that our model is stable. We want to use the coefficient output of the VAR(5) model in order to calculate the companion matrix $A$, which we need to compute a fundamental price estimate. However, this requires that the VAR(5) model is stable. This implies that there must be no unit roots in our data and that the model should ideally be full rank. To formally test these assumptions, we conduct a Johansen trace test and a Johansen eigen test for cointegration. The results of these tests are plotted in table (6.4), where we also report the eigenvalues ($\lambda_r$) for each corresponding rank. In our VAR(5) model, a full rank would imply that five cointegration vectors exist between our five variables. Analysing the output in table (6.4), our results indicate that we can at most reject the null hypothesis of up to 2 cointegrations at 5% significance level in our trace test. Meanwhile, the eigen test rejects a rank of 1 at the same critical level. These tests suggests that our rank is likely 3 or 2, and we can thus not conclude that our VAR(5) model is full rank.

**Table 6.4:** Cointegration Tests of VAR Model Variables.

| Rank $\leq r$ | Eigenvalue ($\lambda_r$) | Trace test | Eigen test |
|---|---|---|---|
| 0 | 0.360 | 127.41*** | 49.57*** |
| 1 | 0.271 | 77.83*** | 35.09** |
| 2 | 0.205 | 42.74** | 25.44* |
| 3 | 0.103 | 17.30 | 12.02 |
| 4 | 0.047 | 5.28 | 5.28 |

* p<0.10, ** p<0.05, *** p<0.01

Table (6.4) displays the results of the Johansen Trace test and the Johansen Eigen test. The eigenvalues are provided alongside critical values of the tests, respectively. The test statistics are calculated on ($n = 116$) observations.

Noting these results, we verify if other tests can indicate whether our model is stable. Chiefly, we must check the modulus of the companion matrix of the coefficients from the VAR(5) model to verify if there are any unit roots in our data. If there are, we would expect to find unit roots larger than one in absolute terms, i.e., outside the unit circle. This would be a major issue, as this would imply that the matrix of the coefficients is not invertible and thus make the companion matrix unstable. Therefore, we test for unit roots in our data using the function *roots()* from the *vars*-package. According to Pfaff, stability in a VAR model implies that all eigenvalues have a modulus of less than one (2008). This is the same as no unit roots with an absolute value more than or equal to one. Table (6.5) below shows the calculated eigenvalues of our VAR(5) model, totalling 25 from our five variable models with five lags. We find that the largest eigenvalue has a value of 0.934. This indicates that our VAR(5) model might indeed be stable and able to be inverted, as needed to compute the fundamental price. However, we do see that the largest modulus is quite large. Therefore, we test the VAR(5) model by varying the number of lags. Testing iteratively with a lag length of 1 up to 8, our VAR model returns eigenvalues of at most 0.96. Therefore, we move on with the analysis under the presumption that our VAR(5) model is stable and able to be inverted.

**Table 6.5:** Eigenvalues of the VAR(5) Model

| | | | | |
|---|---|---|---|---|
| 0.934 | 0.907 | 0.881 | 0.881 | 0.864 |
| 0.864 | 0.831 | 0.831 | 0.808 | 0.808 |
| 0.797 | 0.797 | 0.767 | 0.767 | 0.735 |
| 0.735 | 0.653 | 0.653 | 0.619 | 0.619 |
| 0.615 | 0.615 | 0.518 | 0.518 | 0.376 |

Table (6.5) displays the eigenvalues generated by the VAR(5) model. They are sorted from largest modulus to smallest. An eigenvalue $| \geq 1|$ signifies the model is not stable, i.e., unit roots larger than one.

### 6.2.4 Impact for our Analysis

Comparing our findings for rank and unit roots, we get somewhat conflicting results. The indication that our model might not be full rank is worrisome. At the same time, the eigenvalues tell us that the companion matrix $A$ is likely stable. We opt to move forward with our fundamental price analysis considering these factors. However, to account for the implication of less than full rank, after arriving at a fundamental price, we will estimate a vector error correction (VEC) representation of our VAR(5) model under the assumption of reduced rank. Specifically, we will estimate a VEC under the assumption that rank is equal to two, as that is the lowest level of cointegration we cannot reject with a 5% confidence level in both tests. If the estimated fundamental price is similar to the one predicted by our VAR(5) model, it will give us more confidence in the stability and reliability of the model for further analysis.

## 6.3 Estimating The Fundamental Price

After we have estimated the VAR model and found the tests of the model's stability to be adequate for further analysis, our next step is to compute the fundamental price estimate as outlined in our methodology. As per our equation (4.22), we need an estimate of the companion matrix $A$ in order to rewrite the VAR(5) model into its VAR(1) format. We extract the coefficients from the VAR(5) model and recreate the corresponding companion matrix. The matrix is shown in its entirety in table (A6.1) in the appendix. One advantageous point regarding calculating the fundamental price this way is that it reduces the loss of observations from potentially five observations to one.

After computing the companion matrix $A$, we create the $g1$ and $g2$ vectors with the elasticity parameters $\epsilon_Y$ and $\epsilon_R$. Regarding their value we follow Bergman and Sørensen (2021) and Hott and Monnin (2008) in assuming a base case with $\epsilon_Y = \epsilon_R = 1$. [17] After this, we still need to calibrate the adjustment parameter $\Psi$ to arrive at the most accurate estimation of $s_t$. In order to compute this, we compute an iterative loop where we calculate $s_t$ using equation (4.32) given various $\Psi$ values. Using this method, we find that a $\Psi$ of 0.73 is the optimal value in order to adjust $s_t$ such that average price-to-imputed-rent ratio is as close to the average user cost as possible, in line with our methodology. Having calculated the price-to-imputed-rent ratio, we compute the fundamental price time series estimate as per equation (4.33). Finally, we plot the resulting time series below in figure (6.1).

**Figure 6.1:** Estimate of Fundamental Price

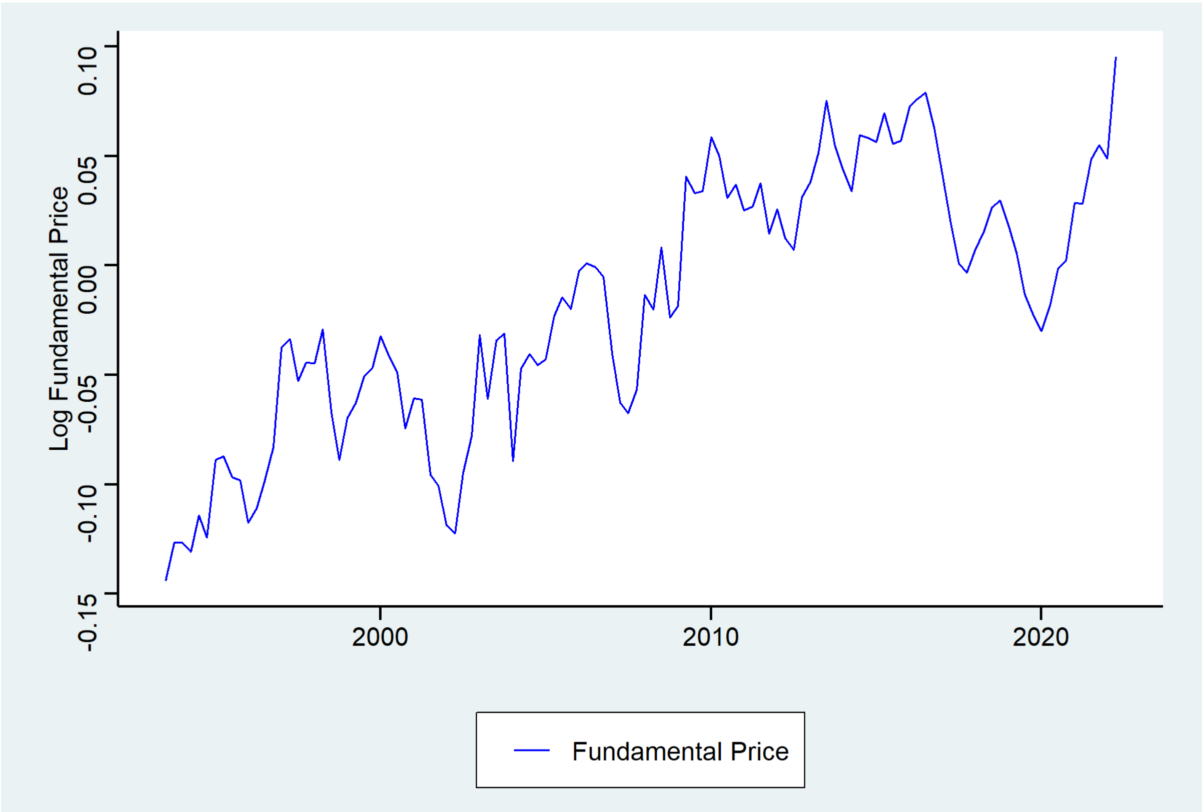


Figure (6.1) displays the estimated fundamental price for the Quarters Q2 1993 - Q2 2022. On the y-axis we have plotted the fundamental price.

[17]We will later test the robustness of these estimates through a sensitivity analysis of $\epsilon_R$.

Naturally, the fundamental price has limited interpretability and use in the analysis without being plotted against the real price. However, as we discussed earlier, when testing the VAR model for stability, we will first compute a vector error correction representation (VEC) under the assumption of $rank = 2$. If this test indicates a sufficiently stable model, despite its lack of stationary, then we continue our analysis and can reasonably draw insights from comparing the fundamental and actual price.

### 6.3.1 A Vector Error Correction Representation

In order to compute the VEC representation, we first utilize the *ca.jo()* function from the *urca*-package, which produces a VECM estimate. Here we compute the VECM under the condition of rank equal to two. We then transform the VECM estimation into a VAR in levels, using the *vec2var()*-function. We now have the coefficients of a VAR model on our data, which have been computed under the assumption of rank being two.

Having acquired the coefficients, the remaining steps in computing the fundamental estimate are the same as with the original VAR(5) model. We calculate the companion matrix in equation (4.23) given the new coefficients and conduct the iterative loop for the price-to-imputed rent ratio. We find that a slightly lower $\Psi$ value of 0.71 is a more appropriate adjustment factor in this case and thus calculate the price-to-imputed rent ratio $s_t$. Using this estimate, we find the VEC representation's estimate of fundamental price. Below in figure (6.2) we plot the VAR and the VEC estimates of fundamental price. The two time series seem to be considerably aligned from a purely visual analysis. The VEC estimate can be said to be slightly more volatile, going beyond the VAR model in most peaks and dips in the data. However, there does not seem to be a considerable deviation.

**Figure 6.2:** VAR and VEC Representation

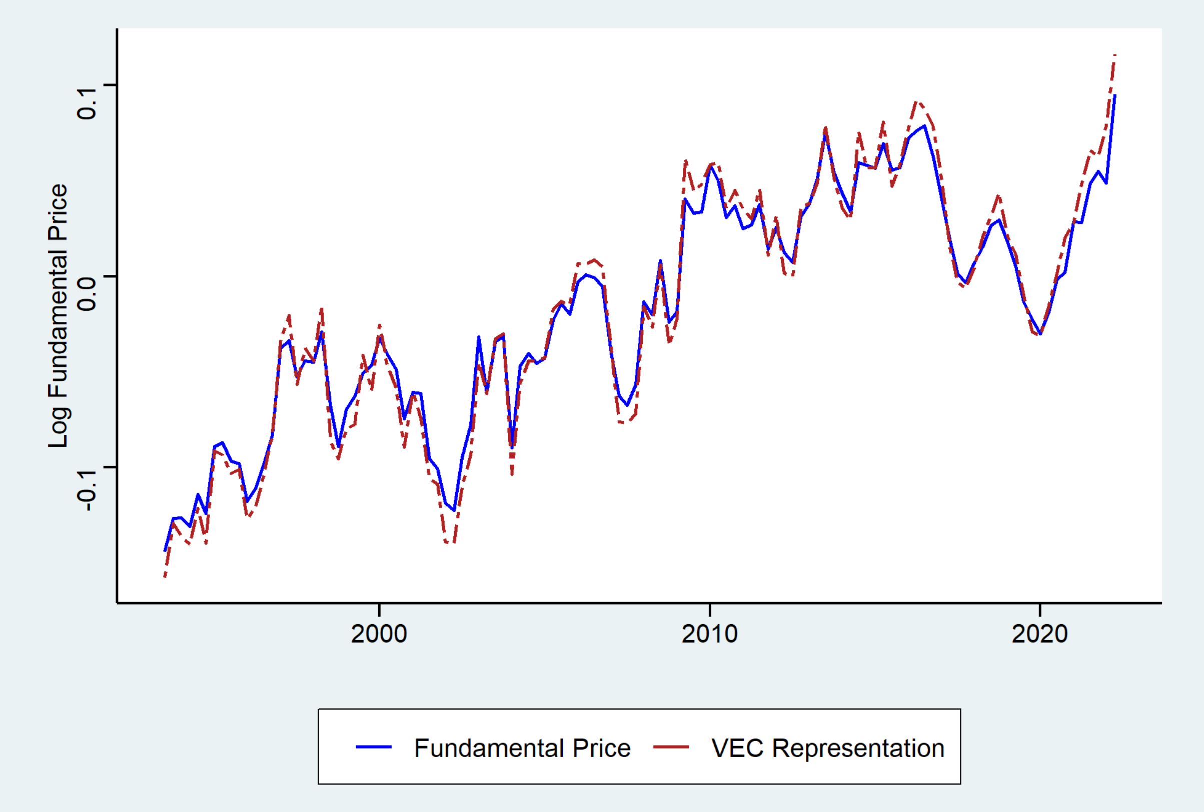


Figure (6.2) displays the estimated fundamental price of the VAR model for the Quarters Q2 1993 - Q2 2022 under the assumption of full rank together with the VEC model estimate under the assumption that rank = 2.

We perform additional quantitative analysis on the two time series estimates to check for any significant differences that may not be apparent in the visual plot. The summary statistics of both series are displayed in Table (6.6). Upon comparing the values, we see that the statistics are consistent with what we can see from Figure (6.2), indicating that the time series have roughly the same mean and standard deviation. Therefore, our VAR model seems to provide a stable estimate of the fundamental price, even when considering the potential non-stationarity of the data. Based on these findings, we can continue with the assumption that the model is stable.

**Table 6.6:** Descriptive Statistics of VAR and VEC

| Variable | Min | Median | Mean | Max | St. Dev |
|---|---|---|---|---|---|
| VAR | -0.14398 | -0.01848 | -0.01808 | 0.09516 | 0.05848 |
| VEC | -0.15937 | -0.01591 | -0.01916 | 0.11870 | 0.06625 |

Table (6.6) displays a comparison of descriptive statistics of the fundamental price calculated by our VAR (5) model and the VEC model representation estimate.

## 6.4 Fundamental Versus Actual House Price

With the fundamental price estimated and evaluated as built on a stable VAR(5) model, we can compare it to the actual house price and analyse the degree to which the Norwegian housing market is in line with what fundamentals would suggest. In figure (6.3) below, we have plotted our fundamental house price against the log of the actual house price. Additionally, we apply 95% confidence bands to the fundamental house price.

From the plot, we can see that the fundamental house price at the start of the sample was decently higher than the actual price levels. This is a reasonable estimate, as we have estimated the fundamental price on a subset of the data beginning at the lowest point after the bursting of the Norwegian housing bubble in the early 90s. We can hypothesize that the real price would have dropped significantly more than the fundamentals underpinning it. The two prices then gradually converge, and the actual price briefly crosses above the fundamental price for the first time in Q4 2000. Then, in Q3 2001, we see a diverging path, as the fundamental price drops significantly, bottoming out in Q2 2002. However, we note that the difference is still comfortably within the limits of the 95% confidence bands of the fundamental price. The two time series then converge again, largely in sync until 2006, when they diverge. Interestingly, this time we see a marked discrepancy in price development. The fundamental price declined sharply, hitting a low in Q3 2007, before recovering swiftly. Meanwhile, the actual price started to rise at a faster growth rate, hitting a peak in the same quarter, before beginning a slow decline until Q4 2008. In this case, we can see that the actual price verges on crossing the upper confidence band at the peak.

**Figure 6.3:** Fundamental and Actual House Prices

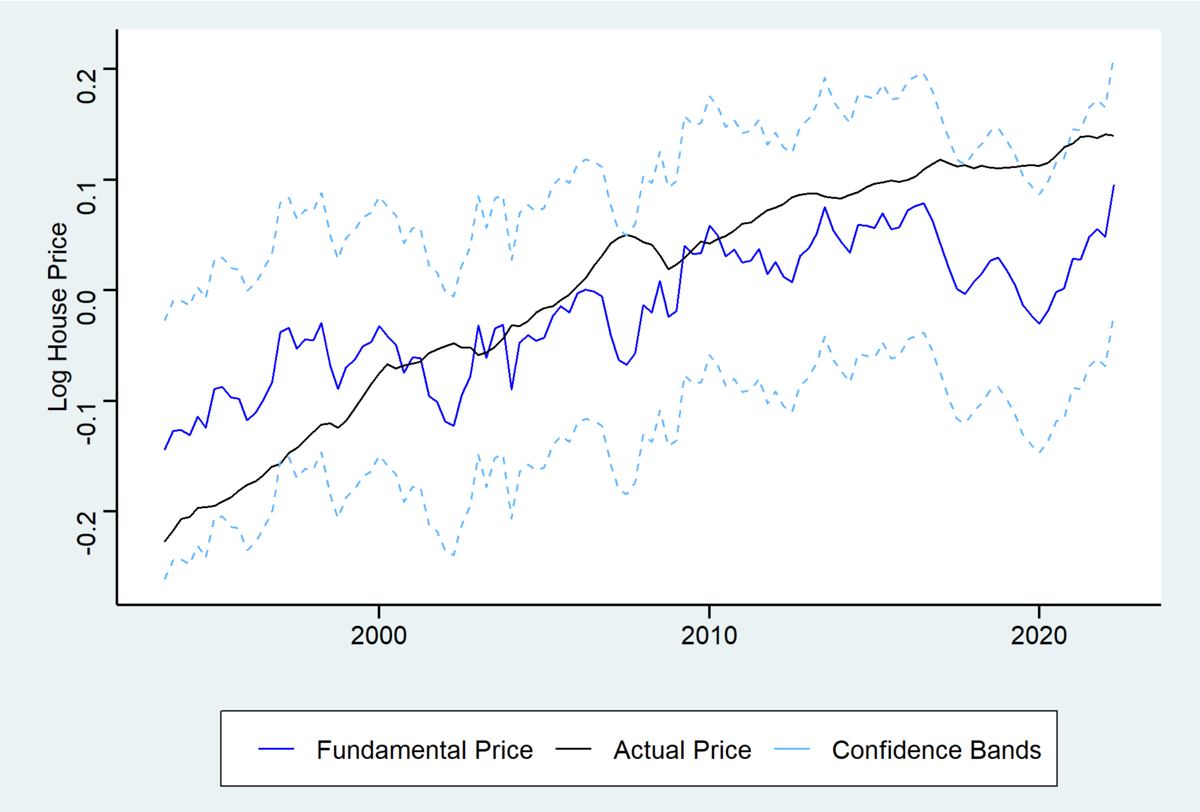


Figure (6.3) displays the estimate of fundamental house prices generated by the VAR model together with the logarithm of the actual house price. The confidence bands are at the 95 % significance level, represented by two standard deviations added or subtracted to the fundamental price.

Analysing changes in the variables utilised to calculate fundamental price, we find that rises in user costs likely drive the decline of fundamental price in both 2001 and 2006. We explore further and compare our user cost variable with the statistics on inflation we received from Statistics Norway, as well as the historical development of the policy rate set by Norges Bank (2022). We find that in the case of the 2001 drop, there were high interest rates and inflation initially. However, when inflation dropped, the interest rates maintained a high level for some time. Meanwhile, in 2006, we find that this period coincides with a sharp increase in the policy rate, increasing from 2.75% in Q2 2006, to 4.5% a year later. We expect that the policy rate will influence Norwegian household mortgage rates, a key component in the user cost. Furthermore, inflation fell significantly in 2007, giving further credence to a significant increase in the real user costs of the period.

Comparing the time series further, we see that the real price was more or less aligned with fundamentals from this point on and until Q3 2016. At that point, the fundamental price hit a peak which would only be surpassed by our last quarter, Q2 2022. From late 2016, the figures indicate that actual and fundamental price started to diverge. The gap reached its greatest size in Q1 2020. The figure indicates that the actual price crossed the upper 95% confidence band for a limited number of quarters. After that, we see a general convergence until the end of the sample. However, it is driven purely by a marked increase in the fundamental price. At the end of the sample period, fundamental price is still considered below actual price levels, but the gap is insignificant at a 95 % confidence level.

Overall, we see that the fundamental price is fairly consistent with the actual price over the sample period, although there are periods of diverging trends. Moreover, we find that the difference between the mean actual price $\bar{p}^a$ and mean fundamental price $\overline{p^f}$ is only $\approx 2\%$. We note that this is similar to the results of Bergman and Sørensen.

### 6.4.1 Testing Sensitivity of the Fundamental Price

Our estimate of fundamental price is based on several important assumptions. Therefore, it is useful to test our estimate's robustness to variations to some of these assumptions. In this part, we will conduct a sensitivity analysis of our price estimate given variations of the long-run imputed rent elasticity of housing demand, $\epsilon_R$, and the adjustment parameter $\Psi$.

#### 6.4.1.1 Sensitivity of Elasticities

When testing the sensitivity of the elasticities, we follow Bergman and Sørensen in keeping $\epsilon_Y$ constant and varying $\epsilon_R$. This is further motivated by Englund (2011) who emphasises that $\epsilon_Y$ is equal to 1 while $\epsilon_r$ is less than 1. Therefore, in figure (6.4) below, we have estimated the fundamental price for the three instances of $\epsilon_R = 1.0, 0.7$ & $0.5$. Looking at the plot we see that there are some changes with a reduced elasticity. The figure suggests that a lower elasticity will return a lower fundamental valuation in the early parts of the data set, until the early 2000s. The estimates are then more or less in line until 2010 when we see that the lower elasticities indicate a consistently higher fundamental estimate until 2016. From then on until the end-of-sample, the estimates are again in line with each other. More generally, we see that the discrepancy is higher with a larger difference in the elasticities.

**Figure 6.4:** Fundamental House Prices With Different Elasticities

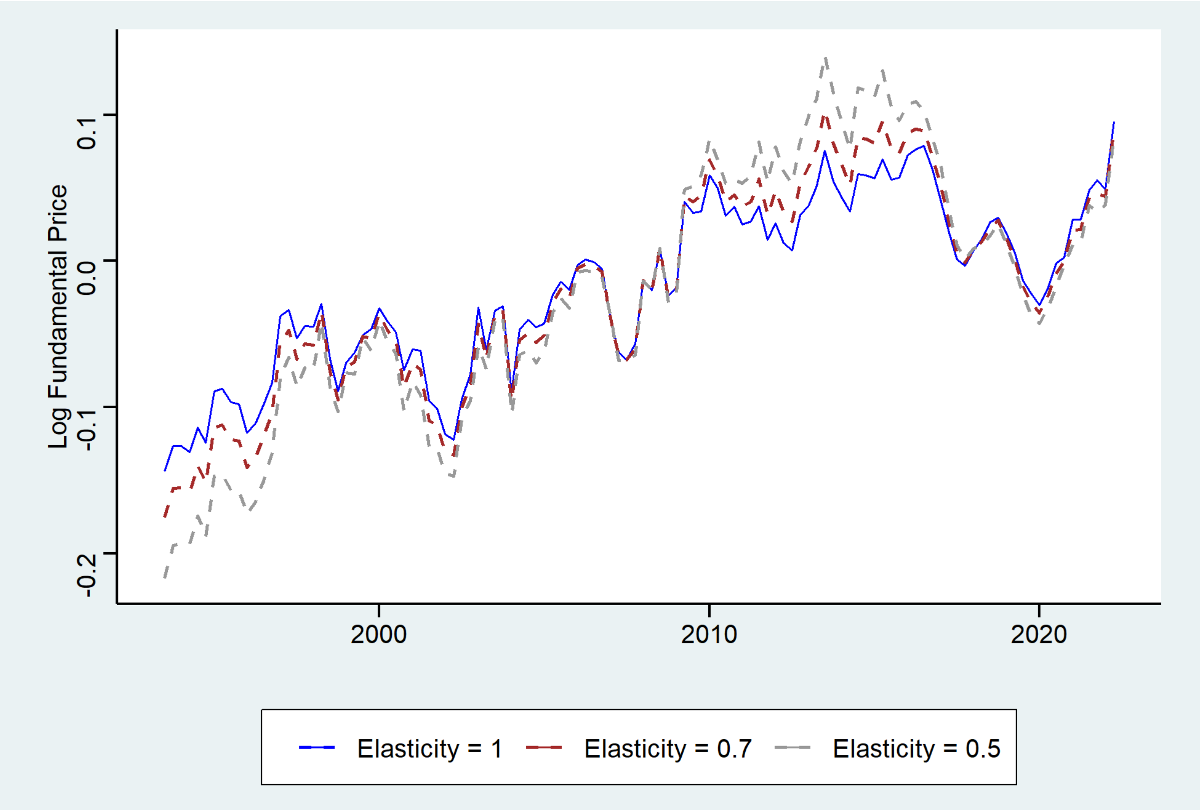


Figure (6.4) displays the fundamental house price and the effect of changes to the elasticity parameter $\epsilon_R$. The test are for sensitivity of it being 1, 0.7, 0.5.

All things considered, we see that plot seems to indicate that there is some effect on the final fundamental price of a variation in $\epsilon_R$. However, the differences are not overly large. Furthermore, the fact that the periods of increased differences return estimates that are both higher and lower than the base case ($\epsilon_R = 1$) suggests that the estimate is not overly sensitive to changes in elasticity, as the mean of the time series is not meaningfully affected. Overall, we presume that the fundamental price estimate is robust to variances of the long-run imputed rent elasticity of housing demand.

#### 6.4.1.2 Sensitivity of Adjustment Parameter Ψ

Next, we test the robustness of our estimate to variations of Ψ. This value must be between 0 and 1, as we showed in the methodology section. We test our chosen Ψ value of 0.73 against the following Ψ values: 0.63, 0.68, 0.78, 0.83, which equal a 0.05 and 0.10 variation around the optimal level. The resulting variations of price estimates are plotted in figure (6.5) below. Analysing the effects of changes in the adjustment parameter, we note that the effect seems to be most notable in the first half of our estimate, with the estimates converging into a somewhat smaller overall plain from then on. However, evaluating the effect more broadly, it does not seem to be large variances nor any meaningful changes to the directions of fundamental price. We therefore conclude that the estimate is fairly robust for changes in Ψ.

**Figure 6.5:** Fundamental House Price and Sensitivity of Ψ Estimate

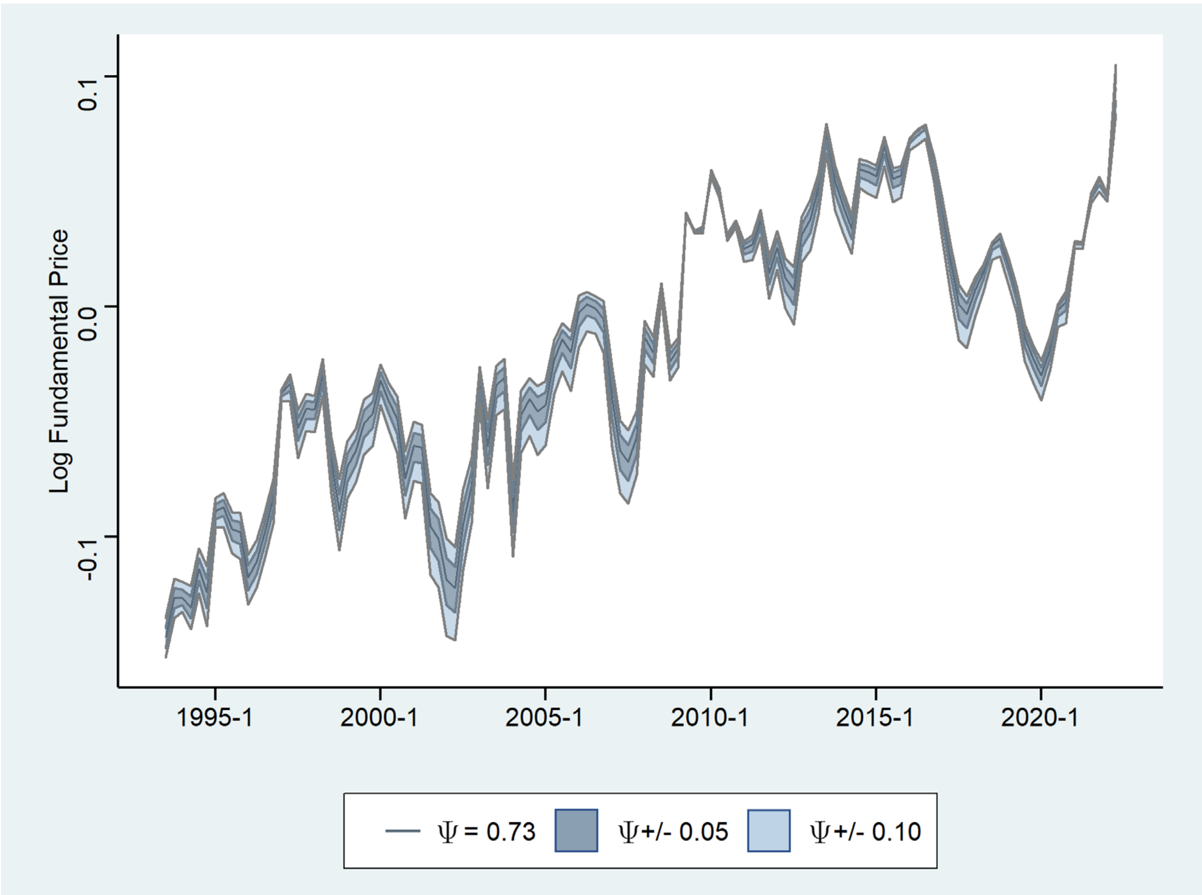


Figure (6.5) displays the fundamental house price and the effect of changes of in Ψ. The sensitivity of this parameter is tested by adding and subtracting 0.05 and 0.10 to our base estimate. The base case Ψ has a value of 0.73.

### 6.4.2 Fundamental House Price and the Effect of Sub-Samples

Our analysis focuses on a subsample from Q2 1993 to Q2 2022. However, as mentioned at the beginning of the analysis, we will now evaluate the fundamental price with the whole sample set, including Q1 1982 to Q1 1993. The choice to limit the scope of the data set was made on the assumption of a potential structural break in the data. Thus, we would expect to see apparent differences in the coefficients and, thus, the estimates. We, therefore, compare our above estimate with an estimate of the fundamental price for the complete sample (Q1 1982 to Q2 2022) and the disregarded sample period (Q1 1982 to Q1 1993).

We construct the two new estimates following the computational method explained earlier in the analysis, with the new input data. The results are plotted below in figure (6.6), together with the actual price over the entire period. As the VAR model has been estimated on scaled data, we subtract the mean of the full actual price and the full sample plot in the period (Q2 1982 to Q1 1993) and for the period (Q3 1993 to Q2 2022).[18] So as to remove doubt on how to interpret the plot, we distinguish between the two demeaned subsamples with a vertical line at the point of junction.

When analysing the graph, we assume that if there was no structural break, then the expectation would be that the full sample estimate returns a somewhat similar estimate in both sub-periods as the period-specific estimates. Thus, comparing the full sample and the subset for 1982-1993, we would expect to see at most minor differences. However, there seems to be a clear deviation between the two estimates, starting at the peak of the housing bubble in the Norwegian market. This could indicate that the data are structurally different, although we note that there is not an overly major difference. Comparing the full sample and our chosen sample from 1993 to 2022, we find further discrepancies. The plots indicate that the inclusion of earlier observations affects the coefficients such that the full sample estimates a higher fundamental price on average than our chosen subset. This further strengthens the argument of the estimated price being sample dependent in our case.

[18]We note that the fundamental price is now estimated as a VAR(1) model, which requires an initial period in order to begin estimation, thus all estimates on the subsets are one period ahead.

**Figure 6.6:** Fundamental House Prices on Different Sub-Samples

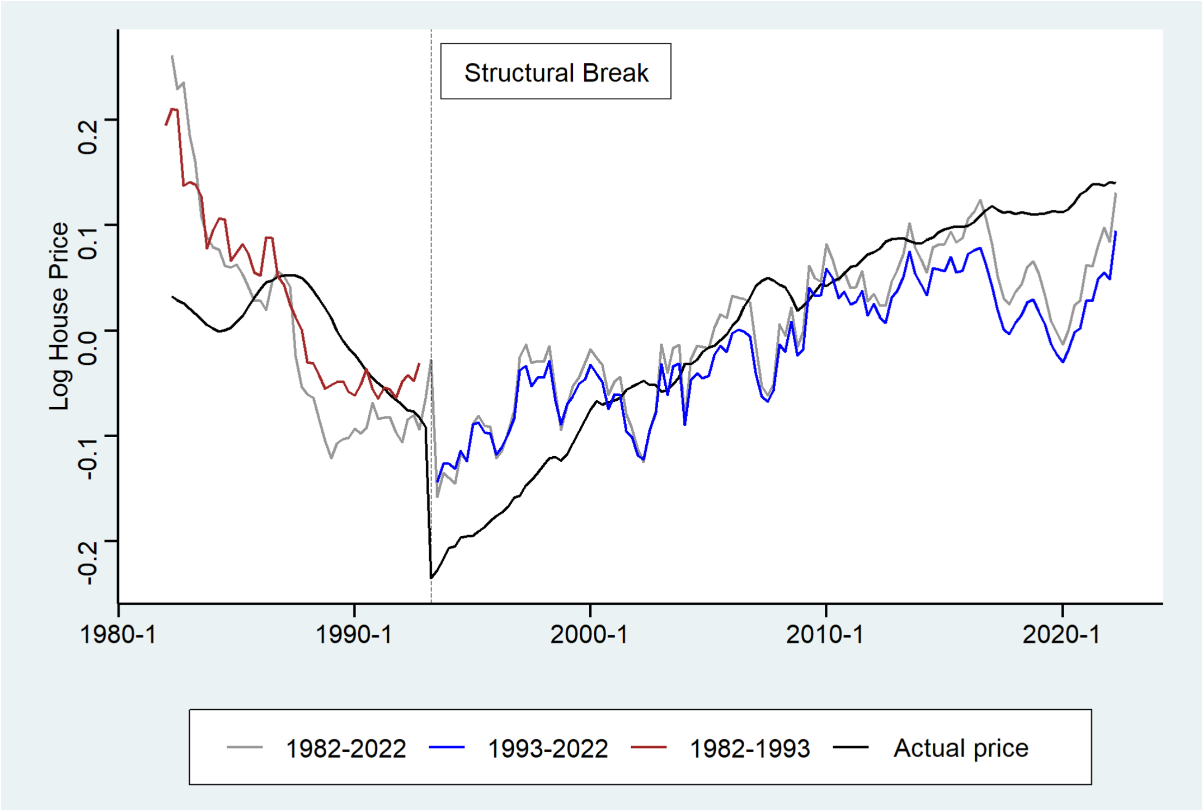


Figure (6.6) displays the fundamental house price and the effect of calculating the estimate on different sub-samples. We here test how sample Q1 1982 - Q1 1993, Q1 1982 - Q2 2022, and Q3 1993 - Q2 2022. The actual price is also plotted for reference for the entire period. All sub-samples are demeaned, so that the individual period has a mean of zero. The full sample uses the estimated $\Psi$ so that we can see to what degree it is useful in predicting the results in the main sample. For the disregarded sample we re-estimate $\Psi$ to get an as precise estimate as possible on the fundamental price of the period.

In addition to an analysis of the fundamental estimations, we investigate other structural differences. Specifically, differences in unit roots can inform us of underlying differences in the subsets. As such we use the function *roots()* to find the eigenvalues of the Q2 1982 - Q1 1993 estimate. We find that the VAR model has eigenvalues larger than one[19]. This indicates that the companion matrix of the coefficients is unstable. Considering these aspects, we find that there seems to be reasonable evidence for our choice of estimating the fundamental price the Q2 1993 - Q2 2022 subset of the data.

[19]The complete set of eigenvalues belonging to this subset are presented in the Appendix table (A8.1)

## 6.5 Analysing the Interaction of Actual and Fundamental House Prices

With the fundamental house price determined, we can analyse the relationship between the fundamental and actual price in the Norwegian housing market, as per our second research question. Concretely, we are looking to evaluate whether there is a cointegration relationship between the fundamental and actual price. We are also interested in assessing the nature of the potential relationship. As outlined in section (4.5.2) in the methodology, we will expect a convergence towards an equilibrium, in the long run, and for the convergence to be purely one-sided.[20] I.e., the actual price is influenced by levels of fundamental price, while the fundamental price is independent of the actual price. In this part of the analysis we aim to analyse both the existence of a potential relationship, and whether theoretical assumptions on the nature of the relationship hold true.

### 6.5.1 Finding the Optimal Lag Length for VECM

In order to analyse the relationship between fundamental and actual price, we must model the relationship between the time series. As laid out in our methodology, we follow Bergman and Sørensen in doing this through estimating a vector error correction model (VECM). Our inputs in the VECM is the fundamental price and the actual house price between 1993 and 2022. We thus combine these time series in a common a vector, $q_t$.[21]

Before we can model a VECM and test the validity of imposing restrictions on the $\alpha$ and $\beta$ vectors however, we must first decide on the lag length of the new model representation. Drawing upon the theoretical framework of VECMs, we can find an estimate of lag length from the underlying bivariate VAR model.[22] Thus, we again make use of the *vars*-package and the function *VARselect()*, which returns four estimates of the optimal lag length for a potential VAR model of $p_t^a$ and $p_t^f$. In table (6.7) we see the suggested lag lengths of the four returned information criteria. We see that three tests out of four indicate a lag length of five, while the SC (BIC) tests indicate an optimal lag length of 2.

[20]In the short- and medium run, however, temporary housing bubbles and other factors of friction might create deviations counteracting this convergence.

[21]The definition of the vector is $q_t \equiv [p_t^a, p_t^f]'$, and was introduced in section (4.5).

[22]As mentioned in section (3.3), a VECM is a VAR model in first differences, $(I(1)$. The changes in $q_t$ per period $(\Delta q_t)$ is therefore the estimated vector.

**Table 6.7:** Test for Optimal Lag Selection of VECM

| Test | AIC(n) | HQ(n) | SC(n) | FPE(n) |
|---|---|---|---|---|
| Lags | 5 | 5 | 2 | 5 |

Table (6.7) displays the four different proposed lag lengths for our VECM. We in this context opt for the lag length indicated by AIC HQ and FPE, and choose 5 as our lag length.

For estimating the lag length of the VECM, we put a high emphasis on the different information criteria. This is because we are interested in testing the theoretical relationship between real and fundamental price. Assessing which lag length to use, we consider that three out of four information criteria return the same lag length, including the AIC. Therefore, although there are advantages to using the SC, we ultimately choose to move forward with a lag length of 5 for the VECM.[23] We conduct tests of the residuals, and find that there is no issue with autocorrelation nor normality in our estimate. The tests do note some issues with heteroscedasticity, but overall we find the lag length as suitable for our analysis.[24]

### 6.5.2 Testing Whether Cointegration Rank Equals One

Now that we have the found the suitable specifications for our VECM, we can evaluate whether there is a cointegration vector between the actual house price, $p_t^a$, and the fundamental house price, $p_t^f$. Using the *ca.jo()*-function we estimate our VECM. Importantly, we clarify that the lag length will be 4, and not 5 as returned by the tests above. This is because the tests were for the underlying VAR in levels. However, when estimating the VECM, lag length is $(p-1)$ of a VAR model.[25]

The results are shown in table (6.8). We find that a rank of at most one cointegration cannot be rejected at a 5% confidence level. Meanwhile the hypothesis of at most 0 cointegrations can be rejected at the 1% level. Therefore, we reasonably assume that we have one cointegration vector in our model.

[23]Schwarz Criterion is considered to have a higher probability of containing the "true" model when estimating a model to fit historical data (Chakrabarti and Ghosh (2011).

[24]We have included the results of these residual tests in table (A4.1), which can be found in the Appendix.

[25]We note that this is also how the *ca.jo()*-function computes the VECM, i.e., the input is $p$.

**Table 6.8:** Cointegration Tests for VECM

| Rank $\leq r$ | Eigenvalue | Trace test |
|---|---|---|
| 0 | 0.179 | 30.64*** |
| 1 | 0.076 | 8.72* |

* p<0.10, ** p<0.05, *** p<0.01

Table (6.8) displays the results of the Johansen Trace test, where the null hypothesis is a rank lower than or equal to the evaluated rank. The eigenvalues are also provided.

We display the loading matrices of the estimated VECM in table (6.9) below. These are the $\alpha$ and $\beta$ matrices mentioned previously. The $\beta$ matrix has been normalized so that the first row is equal to one. The VECM has been estimated with a constant, but no linear trend. We note that this VECM is unrestricted, i.e., $\alpha\beta$ are allowed to freely choose the optimal value.

**Table 6.9:** Estimated $\alpha\beta$Loading Matrices for Unrestricted VECM

| Adjustment Coefficients $\alpha$ | | | |
|---|---|---|---|
| Variable | $p^a.l1$ | $p^f.l1$ | Constant |
| $p^a.l1$ | -0.0201 | 0.0004 | 0.0000 |
| $p^f.l1$ | 0.0041 | 0.0786 | 0.0000 |

| Coefficients of the Cointegration Matrix $\beta$ | | | |
|---|---|---|---|
| Variable | $p^a.l1$ | $p^f.l1$ | Constant |
| $p^a.l1$ | 1.0000 | 1.0000 | 1.0000 |
| $p^f.l1$ | -1.1953 | -2.2709 | 0.5251 |
| Constant | -0.1360 | -0.0199 | 0.0559 |

Table (6.9) displays the estimated $\alpha$ and $\beta$ loading matrices of the unrestricted VECM, where we have allowed for a constant, but not no linear trend. The VECM is estimated on ($n = 116$) number of observations.

### 6.5.3 Imposing Cointegration Restrictions On $\alpha$ and $\beta$

The next issue at hand is imposing the necessary restrictions on our VECM, such that it is consistent with the theory of fundamental valuation. We begin by referring back to table (6.9), where we returned the matrices of the unrestricted VECM. It is these matrices that will now be restricted, so that the relationship can be estimated under the assumption that the relationship is of a specific nature.

Here we follow the method we outlined in our methodology. The restrictions on $\beta$ are imposed on the assumption that there exists one cointegration relationship between the variables such that $\beta = [1, -1]$, meaning the variables are mean-reverting.[26] The

[26] We note that restrictions can be thought of as functionally removing columns of the identity matrix

restrictions on the adjustment coefficient vector $\alpha$, such that $\alpha_1 < 0$ and $\alpha_2 = 0$, imply that the fundamental house price is unaffected by a gap in prices, while the actual house price is assumed to be negatively influenced by a gap between the actual and fundamental house price. [27]

Using the *ablrtest()*-function from the *urca*-package, we implement the chosen restrictions to the $\alpha$ and $\beta$ matrices[28]. The results are shown below in table (6.10), where we plot the test statistic and p-value of the log-likelihood test, as well as the returned eigenvalues. The null hypothesis is that a VECM based on these restrictions is a plausible representation of the relationship between real and fundamental price, given the existence of one cointegration.

**Table 6.10:** Hypothesis and Result of VECM

| Hypothesis | Test stat | p-value | $\tilde{\lambda}_1$ | $\tilde{\lambda}_2$ |
|---|---|---|---|---|
| $\mathcal{H}_{1,1} \mid H_1(r=1)$ | 0.2004 | 0.6544 | 0.1777 | 0.0000 |

Table (6.10) Displays the test statistics, p-value and eigenvalues for the restricted VECM. The null hypothesis $\mathcal{H}_{1,1} \mid H_1(r=1)$ is for these restrictions to return $\alpha$ and $\beta$ values which are significant, i.e., the restricted VECM can plausibly describe the relationship between fundamental and actual price. This is given a cointegration rank of one between the two time series.

The p-value suggests that the hypothesis of the restrictions on $\alpha$ and $\beta$ cannot be rejected. In equation (6.1), we show the returned estimate of the $\alpha$ and $\beta$ vectors of the restricted VECM. Here, we see that $\alpha_1$ takes a value less than one, which fits neatly with the assumption of a negative reaction by actual price to an increase in the gap. Overall, we see that the theoretical assumptions seem to hold.[29]

$$\alpha = \begin{bmatrix} -0.0178 \\ 0.0000 \end{bmatrix}, \; \beta = \begin{bmatrix} 1.0000 \\ -1.0000 \\ -0.1478 \end{bmatrix} \tag{6.1}$$

of $\beta$, as shown by Pfaff in his book "*Analysis of Integrated and Cointegrated Time Series with R* (2008).

[27] $\alpha_1$ and $\alpha_2$ refer to real and fundamental price, respectively, because it refers to the row order of the vector $q_t$.

[28] The matrices of the restrictions are provided in the Appendix in section (A7)

[29] As we have included a constant term estimating the VECM, the cointegration vector $\beta$ also includes the constant -0.1478

### 6.5.4 Evaluating Cointegration and Gap Stationarity Over Time

In addition to determining the significance of the tests for the subset of 1993-2022 as a whole, it is useful to evaluate the test scores over time. By doing such an analysis, we can evaluate to what degree the relationship between the variables is stable throughout the sample. If either the hypothesis of a cointegration relationship or stationarity of the gap can be rejected at different points during the period, then this can indicate a disconnect between the real and fundamental price. We follow Bergman and Sørensen and test for cointegration and restrictions on a smaller sample set. We then recursively add one more observation and recalculate, until we reach the end of the sample.

We begin by dividing our data set into a training and a test data set. We decide to split the sets such that roughly 2/3 of the observations are in the training set, and the remaining 1/3 in the test set.[30] In figure (6.7), we plot the recursively estimated scores of both statistical tests. It is useful to briefly explain what the figure shows. In the trace test plot in (6.7a), we follow Bergman and Sørensen in plotting the test level of the probability for both $r = 0$ and $r \leq 1$. To ease visual interpretability, we divide the 5% critical level score by the test statistics for each test. The result is that if either of the test statistic has a value less than one, then that implies a rejection of that null hypothesis. Correspondingly, a value above one, indicates that we cannot reject the null hypothesis. Meanwhile, figure (6.7b) shows the evolution of the log-likelihood test's p-value of the restricted VECM being a plausible representation of the cointegration relationship, where we have added horizontal lines at the 5 and 10% significance level.

[30]This split corresponds to the period Q3 1993 : Q4 2012 and Q1 2013 : Q2 2022, respectively.

**Figure 6.7:** Recursive Trace Test and LR Test.

(a) Recursive Trace Test (b) Recursive LR plot

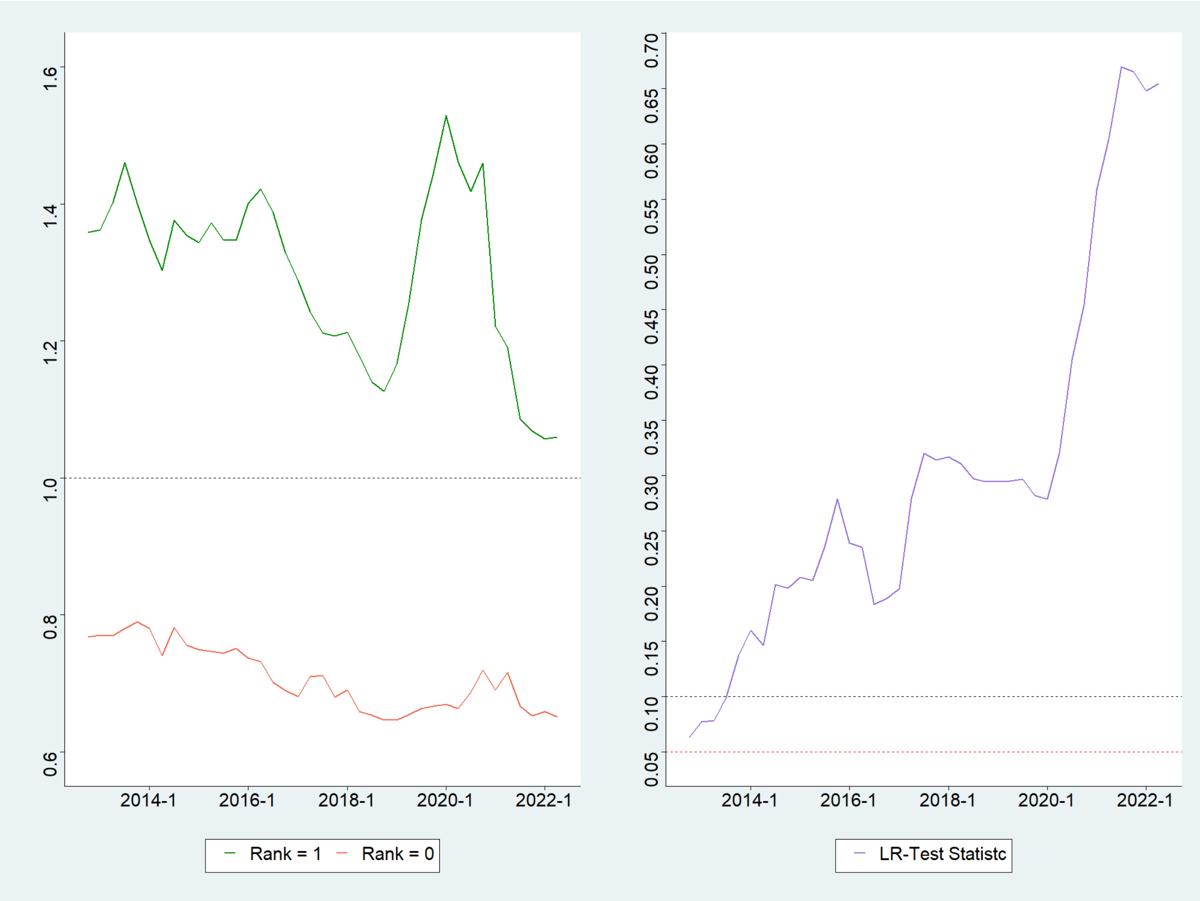


Figure (6.7a) displays the recursive Trace test where we test the null hypotheses that the rank (number of cointegrations) is zero or one. Figure (6.7b) displays the recursive likelihood-ratio test for the null hypothesis that the gap between actual and fundamental price is stationary.

Assessing the results of the trace statistics, we note that the trace test for rank $\leq 1$ is fairly stable. We see that the statistic never crosses below the 5% significance level during the estimated period. There we cannot reject the null hypothesis at any point. Comparing with the test for rank $= 0$, we see that we can always reject the null hypothesis of zero cointegration. At the same time, we see that there are instances where the test for rank $= 1$ is closer to crossing the threshold, particularly during the final observations. There are also some indications that the test statistic is trending towards a drop below the 5% threshold. However, on the whole we find that the tests indicate that the assumption of cointegration is time consistent.

Evaluating the tests of the restricted VECM, we see that the p-values are consistently above significance level of 5%. However, the first few observations indicate a p-value below a 10% significance level, questioning somewhat the validity of the restrictions in the early part of the test data set. At the same time, we see that the p-value rises quickly, reaching a comfortable level above both the 5 and 10% level by 2014 onward. The plot indicates a continues rise in the p-value over the period. This development could be explained by the sample size of the estimation, where each added observation leads to less significance of any observation with sizeable gap in the price levels. Overall, the results suggest that we can not reject the null hypothesis of a VECM based on our restrictions on the $\alpha$ and $\beta$ loading matrices represent a plausible representation of the cointegration relationship over time.

### 6.5.5 Concluding an Analysis of Fundamental and Actual Price

Bergman and Sørensen remark that a disconnect in the housing market can be seen through a combined analysis of the trace test and the LR test of the bivariate VECM. In this instance, for the Norwegian housing market, we see that the trace tests do not indicate a rejection of a cointegration relationship between fundamental and actual price. Additionally, the log-likelihood tests indicate that our restrictions are valid, meaning that the VECM is a plausible representation of the cointegration relationship, under the restrictions that the gap is stationary $I(0)$, and that only actual price is influenced by a gap between the two time series. All things considered, we do not have any empirical reason to suggest that there has been a disconnect between fundamental and actual price in the Norwegian housing market from Q1 2013 to Q2 2022.

# 7 Discussion

In the analysis, we have derived a fundamental price estimate and analysed the development and relationship between the fundamental and actual house price in the Norwegian housing market for the last 29 years. In this section, we will explore the implications of our findings, discuss the validity of our results, and suggest areas of future research beyond the scope of this thesis.

## 7.1 Robustness of Analysis

### 7.1.1 Robustness of Results

Before comparing and discussing our analysis, we briefly summarize our results. Our analysis shows that our VAR model is reasonably stable. Regarding the estimate of fundamental price, we find the actual price to be relatively aligned with what the fundamentals indicate for the examined sample. Regarding our second research question we find that there is a cointegration relationship. Furthermore, we can estimate a plausible representation of the relationship where there is a stationary gap and only actual price reacts to a difference between the prices. An important question is to what degree these findings hold and how robust the fundamental price estimate is.

Evaluating the results, we conducted tests in the analysis to gauge the robustness fundamental price estimate. This is the foundation of the entire analysis. Specifically we tested the sensitivity of fundamental price to differences in the elasticity parameter $\epsilon_R$ and the adjustment parameter $\Psi$. Regarding the elasticity parameter, we find that the fundamental house price estimate is relatively robust to changes in this factor, with minor deviations in fundamental price as the elasticity is changed considerably. The circumstances are similar with the adjustment parameter. These tests indicate a relatively stable result with limited estimated price variance. However, it is likely that larger differences in the parameter level could return quite different estimates of fundamental price. This could be a weakness that our analysis does not account for and could influence the validity of our findings. However, as stated by Bergman and Sørensen, this adjustment coefficient has previously also been subject to arbitrary choice. This history of the parameter value being more or less subjectively considered, indicates that well-founded reasoning for the chosen level of $\Psi$ can be justified.

### 7.1.2 Comparing Our Results to Previous Research

A discussion of the results themselves is most useful as a comparative exercise. We have mentioned Bergman and Sørensen (2021), and Anundsen (2021) frequently in this thesis. Therefore, we base a discussion of our findings primarily on these papers.

Bergman and Sørensen argue that the Swedish housing market has been overpriced since 2014. In our analysis, we can see the same trend in Norway, albeit less pronounced. Anundsen, on the other hand, found the Norwegian house prices to be undervalued until 2016 in his analysis. Anundsen also applied his procedure to Swedish data, finding that the fundamental price was much closer to the actual price than what Bergman and Sørensen found. Comparing our results, we see that although we have applied a method based on Bergman and Sørensen, we find results which are more aligned with Anundsen.

This might be because of several reasons. Firstly, different results might be a consequence of deviations related to the applied methodology and assumptions. Although we have followed Bergman and Sørensen to a large extent, there are deviations. Secondly, there is likely systematic differences between the Norwegian and Swedish data. This can be due to differences in e.g., policies, demographics, or individual features of the respective housing markets. Anundsen on his side analyses both markets, but does not further elaborate on their similarities. This is an apt example of why estimating an "intrinsic" price, such as a fundamental value, is a complex endeavour, as somewhat different methods can yield markedly different results.

Considering the cointegration relationship between the fundamental estimate and the actual house price, Bergman and Sørensen also test for cointegration and stationarity of the gap between actual and fundamental house prices. They find that the Swedish housing market for most of their sample has a cointegration vector, except for when ending the estimation period in 2012. We find similar results. Regarding the likelihood-ratio test and stationarity of the gap, we find somewhat different results from Bergman and Sørensen. Our analysis suggests that the restrictions were valid over most of the sample, but there were indications that the restricted VECM could be rejected early on. Meanwhile, they found opposite results, where the p-value became much smaller towards the end of the sample. Based on these results, it would be interesting to examine what differences in either the underlying data or model there might be. However, on the whole they do not find solid arguments for a disconnect between the Swedish fundamental and actual house prices, which is the same results as ours.

## 7.2 A Discussion on Assumptions

In the process of estimating a fundamental price, we have made a number of assumptions. These range from the choice and accuracy of various variables, to the validity of a VAR or VECM model in the first place. In this section some of these assumption will be discussed, and the potential influence these assumptions could have.

### 7.2.1 Implications of a VAR Model and Time Series Analysis

The choice of model is an important assumption for our thesis, especially related to the different types of VAR models. We chose to follow Bergman and Sørensen in utilising a reduced-form VAR model. As mentioned in the theoretical framework section, the reduced-form VAR version deviates from the other versions as it does not allow for contemporaneous interactions, and only estimates on a basis of prior observed data. If we had chosen another VAR model as e.g., the SVAR, our analysis could have yielded different results. Another issue with our analysis is that we use macroeconomic variables that are integrated of order $I(1)$. Consequently, we difference the variables once in order to have stationary input variables to our model. As a result, we lose one potential observation to estimate the VAR model and therefore, the fundamental price on. In our case one observation is not a significant portion of our data set, but any loss of information is a loss for the analysis.

We also note that this analysis is a time series analysis. While this is a well-documented form of analysis for making in-depth analysis regarding both the past and future, several adverse factors are worth noting. In this regard, one risk is combining aggregated information, indexes and macroeconomic values to make an economic analysis. Using data in different formats increases the chance of measurement errors, or the difficulty in calculating the correct relationship between variables. Furthermore, there is the case of seasonal adjustment. When correcting for seasonality, it also represents a loss of information, as the correction itself is only an estimate of the level of seasonal effect. This can affect the VAR model's accuracy at arriving at the true coefficients.

### 7.2.2 Fundamental Price: a Product of its Input Variables

Our thesis also has had to make assumptions regarding the data and variables. In this part we will address the assumptions and how they might influence our results.

In our thesis we only rely on data which we received directly from Statistics Norway. This is an objective, and well regarded source of information. However, we should note that considering other sources could have led to different outcomes in our analysis, and in certain ways improved the analysis. We could for example considered to utilise only publicly available data published online, in order to increase reproducibility. However, all considered, we feel that the thesis has been well served by utilising this specific data set.

When discussing the variables used in the VAR model, the main issue is to what degree the variable selection is correct. We have assumed that all changes in the fundamental house price are due to changes in the five input variables: actual price, disposable income, user cost, housing stock and rent prices. Assuming that this selection is an exhaustive list of the potential explanatory variables is likely a gross oversimplification of the real world. There are likely other variables that could or should be included in our analysis to make even more accurate analysis.

Furthermore, we have in our analysis had an implicit assumption that an aggregate model is a good description of the outcome on a market where many individuals make decisions. However, this might not be the most applicable approach to provide useful information beyond looking at averages. These issues are also discussed by Bergman and Sørensen, who we base our variable selection on. However, issues like these are ever-present for researchers when estimating or predicting the significance of variables.

**Actual House Prices**

Moving on to a discussion of specific variables, we ought to discuss the nature of house prices. Housing units are not homogeneous commodities, and are an apt example of the issues of aggregation discussed above. In the case of Norway, it is a long and sparsely populated country with lots of different units of housing and local economies. Therefore, our method should only be regarded as a way to better understand the fundamentals of the aggregated housing market. We do not, for example, encourage people to buy a house because our results show that the overall housing market is mildly underpriced. Furthermore, the data on actual house prices does not consider the effect of dividing the property's assets into various pieces. For example, the construction of a new garage, or adding a new bedroom during renovations, would be calculated by Statistics Norway as pure price hike on the property, which will later be put on the market. This represents a further issue with aggregation in the input data.

**User Costs**

Beyond the actual price, we also want to discuss the user costs variable. This specific variable has been shown in our analysis to be of considerable importance in deciding the fundamental price estimate. Assessing the user costs, the main input in calculating the user costs is the data on interest rates. Our interest rate data is, however, based on average levels of a range of interest rates. This thus represents yet another issue of aggregation, and we note that different segments of the market might have access to substantially different interest rates.

Furthermore, user costs consists of an additional three factors, i.e., depreciation rate, property tax and a premium based on risk- and capital constraints. While they were assumed to be constant, this is likely not the case in the real world. One would expect that changes over time in the tax regime might change e.g., the average property tax rate considerably and thus become a more significant aspect of the level of *user cost*. The risk premium might also have changed e.g., if people have become more risk averse after for example the financial crisis. Consequently, our analysis could likely have been more precise and, more importantly, interesting if we had the necessary variable data available for these constants.

## 7.3 Considerations of Fundamental Analysis

We have discussed the results of our analysis, and pointed out some key issues that our thesis faces regarding the various assumptions that we make. However, our thesis necessitates a discussion surrounding the theoretical framework that our method relies on. In this section, we will will begin by discussing the topic of rationality and agents in the market behaving as our economic models assume, before assessing the validity of using the theorem of fundamental asset pricing on housing and in general.

### 7.3.1 Regarding Infinitely Long Time Horizons

An important assumption of our analysis is that agents are assumed to have infinitely long time horizons when buying a unit of housing. This makes it possible to price housing solely by its imputed rent. If, however, their time horizon is not infinitely long, we will need to consider what price the agent can expect to sell the unit of housing for at a given point in the future. A real-life example of agents having an infinitely long time horizon, is that many people expect to scale their housing needs during their lifetime, e.g., when starting a family or as kids start to move out. Thus, assuming infinite horizons is a convenient, but ultimately unrealistic assumption, made by this thesis and by research on the housing market in general. On the other hand, by not making the assumption, we would struggle to compute the agents' behaviour. For example, there is likely not much data available on what consumers expect to sell their house for in the future. One can therefore argue that it is a necessary assumption, however, one should be aware of its limitations.

### 7.3.2 On the Topic of Rationality

In this thesis, we have calculated a fundamental aggregated house price. Theoretically, this will coincide with the price a rational agent would be willing to pay for an average unit of housing. This simplifies modeling behaviour. However, agents are repeatedly found not to act rationally at the aggregated or individual level. While it is convenient to assume agents are rational, it is worthwhile to consider the effect the assumption has on our case. Likely the most troubling assumption tied to rational agents is that they have access to perfect information, and that they can calculate the value of all relevant information based on their preferences. In our analysis of the Norwegian housing market, it might be unrealistic to assume that agents correctly calculate their expectations of future values of rents, prices and user costs to determine their willingness to pay for housing. The discussion on whether agents at the aggregate level behave more or less rationally, however, is a discussion which is still heated in academia.

### 7.3.3 Limitations of Fundamental Valuation

Another important discussion is about the limitations and assumptions of fundamental valuation. The dividend discount model, which our methodology is based upon, is a method for pricing assets based on their expected stream of income. While this model has certain theoretical advantages, it should be addressed that it is based on a number of assumptions that can limit its usefulness. For example, the model relies on assumptions about the rate of return, dividend ratio, and tax and growth rates. These assumptions can make the model more complex and difficult to interpret. In the context of the housing market, the fundamental valuation model may furthermore be less effective due to the difficulty of obtaining reliable and comprehensive data on the market and expectations about the future. Additionally, other theoretical approaches, such as the efficient market hypothesis, state that fundamental valuation is of little use. This hypothesis suggests that all relevant information about a house is reflected in its price, and therefore fundamental analysis may not provide any additional value (Fama, 1998). Despite these limitations, the fundamental valuation model is still widely used in the housing market. One should, however, note that it is important to be aware of the assumptions and limitations of the model and use it with caution.

## 7.4 Application for Future Research

We have so far focused our discussion backwards, i.e., on what we have found and the validity of those results. Looking forwards, we see several interesting topics of future research.

An interesting extension of our work on Norwegian data would be to make it more applicable in describing the actual world. This could be done by making the analysis more specific in terms of, e.g., geographical markets, and types of housing. In this way, one could use our methods to compute an estimate of the fundamental price of the Oslo apartment market. This would help shed new light on the nature of the actual housing market, as experienced by real-world individuals. Of course this requires access to specific data on only Oslo apartments in order to conduct such an analysis, something that might not be easily accessible.

Another interesting topic for research, would be to investigate whether socioeconomic status has an effect on the behaviour in the housing market. With access to the appropriate data, one could perhaps divide population into different segments based on the level of e.g., disposable income. Alternatively, one could approximate the socioeconomic status through classification based on the price of housing, adjusted for geographical area and prevalence of different socioeconomic classes. Obtaining the necessary data for this type of research may be a challenge. However, if these obstacles can be overcome, the results of the research could be valuable in informing government policy and assessing the effectiveness of enacted policies.

A last interesting extension of our research would have been to use our data as a basis for forecasting. Bergman and Sørensen suggests that the model can be used to estimate the effect of policy changes going forward. They specifically test for a situation where paid interest deduction is removed, and evaluate the effect this policy change would have for the Swedish housing market. They do this by feeding an exogenous time line for their user cost variable, and then re-estimating their fundamental price. Such an analysis could also be of interest in a Norwegian context.

# 8 Conclusion

The purpose of this thesis has been to provide new insights into the Norwegian housing market. We chose to do so through an approach based in the theory of fundamental valuation, and the methods introduced by Bergman and Sørensen (2021). In doing so, we sought to answer two specific research questions.

Firstly, we asked to what degree the Norwegian house prices align with what fundamental factors would suggest. On this topic, we estimated a five variable VAR model consisting of housing stock, disposable income, user costs, rent prices and actual prices. We applied the results and estimated a fundamental price of the Norwegian housing market in the period of 1993 to 2022. The analysis shows that the two time series are comparable, but there are instances where they diverge. However, except for a short time frame around 2020, the difference is not significant on a 95% confidence level. Overall, we therefore find that the actual price is largely aligned with fundamentals suggest.

Secondly, we asked whether there is a relation between fundamental and actual price, and if there is, is the relationship in line with what the theory of fundamental valuation would predict? To answer these questions, we applied the computed fundamental price and estimated a VECM based on the two time series of prices. Furthermore, we checked if our results would be time consistent, evaluating the period from 2013 to 2022. We find that that there consistently is one cointegration vector between the actual and fundamental price, indicating a stable, long-term relationship. Next, we imposed restrictions to our VECM in order to test hypotheses of a stationary gap ($I(0)$) and that only actual price adjusts to a difference between them. We then conducted a likelihood-ratio test to determine whether this restricted model is a plausible representation of the cointegration relationship. We evaluated whether these results too, would be time consistent. We consistently find p-values above 5% of our restricted model. We therefore cannot reject the assumptions, indicating that the relationship between the actual and fundamental price is consistent with the theory of fundamental valuation.

# References


ABS (2022). Time series analysis: The basics. Retrieved 11.12.2022 from abs.gov.au: https://www.abs.gov.au/websitedbs/d3310114.nsf/home/time+series+analysis:+the+basics.

Ambrose, B. W., Eichholtz, P., and Lindenthal, T. (2013). House prices and fundamentals: 355 years of evidence. *Journal of Money, Credit and Banking*, 45(2-3):477–491.

Antoch, J., Hanousek, J., Horváth, L., Hušková, M., and Wang, S. (2019). Structural breaks in panel data: Large number of panels and short length time series. *Econometric Reviews*, 38(7):828–855.

Anundsen, A. (2019). Detecting imbalances in house prices:what goes up must come down?*. *Scand. J. of Economics 121(4)*.

Anundsen, A. (2021). House price bubbles in nordic countries? *Nordic Economic Policy Review 2021*.

Bank, N. (2022). Endringer i styringsrenten. Retrieved 06.12.2022 from Norges Bank: https://www.norges-bank.no/tema/pengepolitikk/Styringsrenten/Styringsrenten-Oversikt-over-rentemoter-og-endringer-i-styringsrenten-/.

Bergman, U. M. and Sørensen, P. B. (2021). The interaction of actual and fundamental house prices: A general model with an application to sweden. *Journal of Housing Economics*.

Bergsholm, E. (2022). Eiendomsskattenes historie i norge. *Kart og Plan*, 115(3):303–319. Retrieved 20.09.2022 from Idunn: https://www.idunn.no/doi/abs/10.18261/kp.115.3.8.

Blanchard, O. and Fischer, S. (1989). *Lectures on macroeconomics*. MIT press.

Breusch, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian economic papers*, 17(31):334–355.

Campbell, J. Y. and Ammer, J. (1993). What moves the stock and bond markets? a variance decomposition for long-term asset returns. *The journal of finance*, 48(1):3–37.

Campbell, J. Y. and Shiller, R. J. (1988a). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1(3):195–228.

Campbell, J. Y. and Shiller, R. J. (1988b). Stock prices, earnings, and expected dividends. *the Journal of Finance*, 43(3):661–676.

Campbell, S. D., Davis, M. A., Gallin, J., and Martin, R. F. (2009). What moves housing markets: A variance decomposition of the rent–price ratio. *Journal of Urban Economics*, 66(2):90–102. Retrieved 10.11.2022 from Science Direct: https://www.sciencedirect.com/science/article/pii/S0094119009000400.

Chakrabarti, A. and Ghosh, J. K. (2011). Aic, bic and recent advances in model selection. 7:583–605. Retrieved 07.12.2022 from Sciencedirect: https://www.sciencedirect.com/science/article/pii/B9780444518620500186.

Eitrheim, Klovland, J. T., Qvigstad, J. F., Holter, J. P., Grytten, O. H., Erlandsen, S. K., and Gerdrup, K. (2004). Historical monetary statistics for norway 1819–2003. pages 16–28.

Enders, W. (2008). *Applied econometric time series.* John Wiley & Sons.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007.

Englund, P. (2011). Swedish house prices in an international perspective. *The Riksbank's inquiry into risks in the Swedish housing market.*

Engsted, T., Pedersen, T. Q., and Tanggaard, C. (2012). Pitfalls in var based return decompositions: A clarification. *Journal of Banking & Finance*, 36(5):1255–1265.

Escobari, D., Damianov, D. S., and Bello, A. (2015). A time series test to identify housing bubbles. *Journal of Economics and Finance*, 39(1):136–152.

Eurostat (2022). "distribution of population by tenure status, type of household and income group - eu-silc survey". Retrieved 03.10.2022 from eurostat: https://ec.europa.eu/eurostat/web/main/data/database.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of financial economics*, 49(3):283–306.

Fischer, H. (2011). *A history of the central limit theorem: from classical to modern probability theory.* Springer.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, pages 424–438.

Hailegiorgis, B. A., Belay, K., and Bekele, H. (2011). A time series analysis of structural break time in the macroeconomic variables in ethiopia. *African Journal of Agricultural Research*, 6(2):392–400.

Hänninen, R. H. (1998). The law of one price in united kingdom soft sawnwood imports–a cointegration approach. *Forest Science*, 44(1):17–23.

Hill, R. C., Griffiths, W. E., and Lim, G. C. (2018). *Principles of econometrics.* John Wiley & Sons.

Hott, C. and Monnin, P. (2008). Fundamental real estate prices: an empirical estimation with international data. *The Journal of Real Estate Finance and Economics*, 36(4):427–450.

Jarque, C. M. and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, 6(3):255–259.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, 12(2-3):231–254.

Johansen, S. (1991). The power function of the likelihood ratio test for cointegration. pages 323–335.

Kotze, K. (2022). Vector autoregression models. Retrieved 03.10.2022 from github.com: https://kevinkotze.github.io/ts-7-var/.

Lütkepohl, H. (2007). *New introduction to multiple time series analysis.* Springer Science & Business Media.

Malpezzi, S. (1999). A simple error correction model of house prices. *Journal of Housing Economics*, 8(1):27–62.

Moe, T. G., Solheim, J. A., and Vale, B. (2004). The norwegian banking crisis. pages 1–68.

Palma, W. (2016). *Time series analysis*. John Wiley & Sons.

Pfaff, B. (2008). *Analysis of Integrated and Cointegrated Time Series With R*. Springer, New York, second edition. Retrieved 20.10.2022 from Pfaffikus: http://www.pfaffikus.de.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, pages 1–48.

SSB (2022a). *Consumer price index*. Retrieved 02.10.2022 from Statistics Norway: https://www.ssb.no/en/priser-og-prisindekser/konsumpriser/statistikk/konsumprisindeksen/.

SSB (2022b). Housing price index. Retrieved 13.12.2022 from Statistics Norway: https://https://www.ssb.no/statbank/table/07221/tableViewLayout1/.

SSB (2022c). Makroøkonomisk modell for norsk økonomi (kvarts). Retrieved 10.10.2022 from Statistics Norway: https://www.ssb.no/forskning/makrookonomi/makrookonomisk-modell-for-norsk-okonomi-kvarts.

Stiglitz, J. E. (1990). Symposium on bubbles. *Journal of economic perspectives*, 4(2):13–18.

Stock, J. H. and Watson, M. W. (2001). Vector autoregressions. *Journal of Economic perspectives*, 15(4):101–115.

# Appendix

## A1 Introduction of Variables

Throughout this thesis we have introduced and defined several variables. To ease understanding of the used variables, this subsection will provide an overview of introduced mathematical notation and defined variables.

**Table A1.1:** Mathematical Definitions

| Variable | Mathematical Definition |
|---|---|
| $R_t^H$ | $[i_t(1-\tau_t^i) - \frac{E_t[CPI_{t+1}]-CPI_t}{CPI_t} + \tau + \delta + \eta]P_t - E_t[P_{t+1}] - P_t$ |
| $\gamma_t$ | $i_t(1-\tau_t^i) - \frac{E_t[CPI_{t+1}]-CPI_t}{CPI_t} + \tau_t + \delta + \eta$ |
| $S_t$ | $P_t/R_t^H$ |
| $\overline{m}$ | $\overline{s}^e + \Delta\overline{r}^{He}$ |
| $\Delta r_{t+j}^H$ | $(\frac{\epsilon_Y}{\epsilon_R})\Delta y_{t+j} - (\frac{1}{\epsilon_R})\Delta h_{t+j}$ |
| $b_t$ | $\begin{bmatrix}\Delta p_t^a \\ \Delta r_t \\ \Delta\gamma_t \\ \Delta y_t \\ \Delta h_t\end{bmatrix}$ |
| $\Psi$ | $\Psi = log(1+\overline{\gamma})$ |
| $z_t$ | $\begin{bmatrix} b_t \\ b_{t-1} \\ . \\ . \\ b_{t-n+1}\end{bmatrix}$ |
| $E_t[z_{t+i}]$ | $A^i z_t$ |
| $s_t$ | $c + \Sigma_{j=1}^{\infty}\Psi E_t[\frac{\epsilon_Y}{\epsilon_R}\Delta y_{t+j} - \frac{-1}{\epsilon_R}\Delta h_{t+j} - \gamma_{t+j}] - \gamma_t$ |
| $\hat{p}_t^f$ | $s_t + ((\frac{\epsilon_Y}{\epsilon_R})y_t - (\frac{1}{\epsilon_R})h_t)$ |
| $c$ | $\frac{k}{1-\Psi}$ |

**Table A1.2:** Definitions of Variables

| Variable | Definitions |
|---|---|
| $P_t$ | Real house price of a unit of owner-occupied housing |
| $p_t$ | Logarithm of $P_t$ |
| $\gamma_t$ | User cost of housing |
| $R^H$ | Imputed rent of housing |
| $r^H$ | Logarithm of $R^H$ |
| $H_t$ | Housing stock |
| $h_t$ | Logarithm of housing stock |
| $p_t^f$ | Logarithm of fundamental price |
| $S_t$ | Price-to-imputed-rent ratio |
| $s_t$ | Logarithm of $S_t$ |
| $\epsilon_R$ | long-run imputed rent elasticity of housing demand |
| $\epsilon_Y$ | long-run income elasticity of housing demand |
| $i$ | Nominal interest rate |
| $\tau^i$ | Capital income tax |
| $\tau$ | Effective property tax |
| $\eta$ | User cost premium for risk and credit constraints |
| $\delta$ | Rate of depreciation of the housing stock |
| $\Psi$ | Adjustment parameter |
| $\Phi$ | Coeffisient of lag effect |
| $A$ | Companion matrix |
| $\Delta$ | Indication of First Difference |

# A2 Further Details on Estimation of the Fundamental House Price

In this part of the appendix we show the steps necessary for going from (4.11) to (4.12) in the methodology section.

$$s_t = ln(1 + exp(s^e_{t+1} + \Delta r^{He}_{t+1})) - ln(1 + \gamma_t) \tag{.1}$$

Following this equation, we now define a new variable $\overline{m}$ to be equal to the sum of the average of the logarithm of price-to-imputed-rent ratio ($\overline{s}^e$) and the average period-by-period change in the logarithm of the imputed rent ($\Delta\overline{r}^{He}$). It is thus defined as follows:

$$\overline{m} \equiv \overline{s}^e + \Delta\overline{r}^{He} \tag{.2}$$

Before moving further, we need to introduce the concept of Taylor approximation. This is a necessary step in defining the price-to-imputed-rent ratio.

## A2.1 Taylor Approximation of Two Variables

In general terms, we can define the first-order Taylor approximation formula as follows:

$$f(x) = f(x_0) + f'(x_0) + (x - x_0) \tag{.3}$$

In our situation however, we have two variables which have to be estimated, $\overline{m} + \Delta\overline{r}^{He}$ and $\gamma_t$. Therefore,we have to complete Taylor approximations around both variables. Bergman and Sørensen 2021 introduced $\overline{m} = \overline{s}^e + \Delta\overline{r}^{He}$ and $\gamma_t = 0$ as their approximations. We follow their approach, and get an approximation of:

$$s_t = f(s_1) + f'(s_1)(s - s_1) - (g(s_2) + (g'(s)_2)(s - s_2)) \tag{.4}$$

We first approximate $\overline{s}^e + \Delta\overline{r}^{He}$and use the general Taylor expression with $s_1$ and thereafter include the actual values for $s$ and $s_1$:

$$(ln(1 + exp(s_1)) + (\frac{exp(s_1)}{1 + exp(s_1)}) * (s - s_1) \tag{.5}$$

Which can be rewritten as:

$$(ln(1 + exp(\overline{m}))) + \frac{exp(\overline{m})}{1 + exp(\overline{m})} * (s^e_{t+1} + \Delta\overline{r}^{He} - \overline{m}) \tag{.6}$$

For the second part of the approximation we use $\gamma_t = 0$ and use the general Taylor expression with $s_1$ and thereafter include the actual values for $s$ and $s_2$:

$$(ln(1 + s_2 + \frac{1}{1 + s_2}(s - s_2) \tag{.7}$$

Which can be rearranged to:

$$(ln(1 + 0) + \frac{1}{1 + 0}(\gamma_t - 0)) \tag{.8}$$

This can further be combined. Here we also introduce an adjustment parameter, $\Psi$, which will be explained below. Using it, we get the following:

$$s_t \approx ln(1 + exp(\overline{m})) + \Psi(s^e_{t+1} + \Delta r^{He}_{t+1} - \overline{m}) - \gamma_t \tag{.9}$$

At this point it is convenient to introduce new notation. $\overline{m}$ is defined as follows:

$$\overline{m} \equiv \overline{s}^e + \Delta\overline{r}^{He} \tag{.10}$$

Which is equivalent to:

$$s_t \approx (ln(1 + exp(\overline{m})) - \Psi * \overline{m} + \Psi(s^e_{t+1} + \Delta r^{He}_{t+1} - \overline{m}) - \gamma_t \tag{.11}$$

## A2.2 Further Details on Estimation of the Fundamental House Price

At this point we can follow Bergman and Sørensen (2021) who define that:

$$ln(1 + exp(\overline{m})) = -ln(1 - \Psi) \quad (.12)$$

and that:

$$ln(\Psi) = \overline{m} - ln(1 + exp(\overline{m}) = \overline{m} + ln(1 - \Psi) \quad (.13)$$

Which in turn means that:

$$\overline{m} = ln(\Psi) - ln(1 - \Psi) \quad (.14)$$

Our expression can then be rewritten as:

$$ln(1 + exp(\overline{m})) - (\frac{exp(\overline{m})}{1 + exp(\overline{m})})\overline{m} = -\Psi(ln(\Psi) - ln(1 - \Psi) \quad (.15)$$

And since we know that:

$$ln(1 + exp(\overline{m})) = -ln(1 - \Psi) \quad (.16)$$

We get:

$$-ln(1 - \Psi) - \Psi(ln(\Psi) + \Psi ln(1 - \Psi) \quad (.17)$$

$$-\Psi ln(\Psi) - (1 - \Psi)ln(1 - \Psi)) \quad (.18)$$

Following Bergman and Sørensen (2021)$k$, which is determined as follows:

$$k \equiv -\Psi ln\Psi(1 - \Psi)ln(1 - \Psi) \quad (.19)$$

This takes us back to .11 which we now can rewrite as:

$$s_t = k + \Psi(s^e_{t+1} + \Delta r^{He}_{t+1}) - \gamma_t \quad (.20)$$

This thesis assumes that agents are rational and forward looking. Using the principle of forward iteration we get the following expression of the price-to-imputed-rent of:

$$s_t = c + \Sigma_{j=1}^{\infty} \Psi^j E_t[\Delta r_{t+j}^H - \gamma_{t+j}] - \gamma_t \qquad (.21)$$

Where $c$ is defined as:

$$c \equiv \frac{k}{1 - \Psi} \qquad (.22)$$

## A2.3 Simplified VAR Model Notation

To simplify the notation regarding lag operators in the VAR model matrix representation, we can move the $\Phi$ over and isolate the error term. Then the equation (4.19) becomes:

$$b_t - \Phi_1 b_{t-1} + \Phi_2 b_{t-2} + ... + \Phi_n b_{t-n} = \qquad (.23)$$

We also need to introduce the concept of lag operators to implement the VAR model. This operator can replace the $b_{t-n}$ expression letting us use the general notation as $b_L^2 = b_{t-2}$ and $b_t L^n = b_{t-n}$. In our model, we use this and simplify .23 to:

$$b_t - \Phi_1 L^1 b_t + \Phi_2 L^2 b_t + ... + \Phi_n L^n b_t = \epsilon_t \qquad (.24)$$

Which in turn can be rewritten as:

$$(I_5 - \Phi_1 L^1 - ... - \Phi_n L^n) b_t = \epsilon_t \qquad (.25)$$

Above, since $b_t$ is a vector, we use an identity matrix to recreate the vector. A common way to shorten notation for a series of lagged $\Phi$ values in a time series is to introduce the simplified and convenient notation.

$$\Phi(L) b_t = \epsilon_t \qquad (.26)$$

From this state, we follow Bergman and Sørensen, and define $\Phi(L)$ as

$$\Phi(L) = I_5 - \Sigma_{j=1}^{n} A_j L^j \qquad (.27)$$

Which we, in our case, can rewrite as:

$$(I_5 - \Sigma_{j=1}^{n} A_j L^j) b_t = \epsilon_t \quad (.28)$$

The result is precise notation on how our VAR model utilises the lag operator.

# A3 Data Desription

In the thesis we utilise several data variables. Below we provide descriptive statistics on the variables in their original format, as received by Statistics Norway. Below in table (A3.1) we show information about the data set we received.

**Table A3.1:** Information About Original Variables in Data Set

| VARIABLE | NOMINAL/REAL | INDEX/OBSERVED VALUE | BASE YEAR |
|---|---|---|---|
| HOUSE PRICE | NOMINAL | INDEX | 2019 = 1 |
| RENT PRICE | NOMINAL | INDEX | 2015 = 100 |
| HOUSE STOCK | REAL | OBSERVED VALUE | |
| DISP INCOME | REAL | OBSERVED VALUE | |
| REALRES | REAL | OBSERVED VALUE | |
| POPULATION | | OBSERVED VALUE | |
| CPI | | INDEX | 2019 = 1 |

Table A3.1 contains summary information about original variables in our data set. The table displays if the data is gathered in a nominal or real format, if it is indexed or observed values, and for the indexes, what year is the base year. The REALRES variable is the real interest rate after tax.

Moreover, below in table (A3.2), we provide summary statistics for the original variables in our data set.

**Table A3.2:** Summary Statistics for All Original Values in Data Set

| VARIABLE | MEAN | ST.DEV | MIN | MAX |
|---|---|---|---|---|
| HOUSE PRICE | 0.479273866 | 0.319448235 | 0.134270608 | 1.247431472 |
| RENT PRICE | 71.5372428 | 24.03146241 | 28.23333333 | 112.6333333 |
| HOUSING STOCK | 2872617.694 | 765557.7966 | 1716484.714 | 4442104 |
| DISP INCOME | 260913.9772 | 78549.40959 | 152975.2016 | 416700.1041 |
| REALRES | 0.017832989 | 0.022015712 | -0.037377707 | 0.069426458 |
| POPULATION | 4638869 | | 4107063 | 5440426 |
| *N* | 162 | 162 | 162 | 162 |

Table A3.2 contains summary statistics for the original variables in our data set. The table displays the mean, standard deviation, minimum and maximum value for all the original variables. The REALRES variable is the real interest rate after tax.

# A4 VECM Residual Tests for Autocorrelation, Heteroscedasticity and Normality

In the process of estimating our VECM we tested the statistical properties of our model, given our assumed lag length of five lags. The results of these tests are provided here.

**Table A4.1:** VECM Residual Tests for Autocorrelation, Heteroscedasticity and Normality

| | Autocorrelation | Heteroscedasticity | Normality | | |
|---|---|---|---|---|---|
| Test | LM(5) | ARCH(5) | Jarcue-Bera | Skewness | Kurtosis |
| p-value | 0.1491 | 0.0247 | 0.3626 | 0.2086 | 0.5489 |

Table A4.1 displays the results of our Breusch-Godfrey test of autocorrelation, the ARCH-test of heteroscedasticity, the Jarque-Bera test for normality, as well as multivariate tests for skewness and kurtosis.

# A5 Granger Causality Test of the VECM Variables

As a part of the analysis, we apply the Granger test for causality on the time series of actual and fundamental price. Below in table (A5.1) we test and see that both variables Granger-cause each other at a 5 % significance level.

**Table A5.1:** Granger Cauality Tests for VECM Variables

| | Variable | p-value |
|---|---|---|
| $p_t$ | Fundamental Price | 0.02049** |
| $p_t^a$ | Actual Price | 0.04625** |

* p<0.10, ** p<0.05, *** p<0.01

Table A5.1 displays the Granger causality tests with the null hypothesis that the specific variable does not Granger-cause the other variables that go into our VECM.

# A6 Companion Matrix of VAR(5) Model

In table (A6.1) on the next page we have plotted all factors in the companion matrix. It is a (*kp x kp*)-matrix. In our case we have $k = 5$ and $p = 5$, and we thus have a (*25 x 25*)-matrix, where each group of five columns represent one lag length. Meanwhile, looking at the rows, the first five rows corresponds to the five variables. A characteristic of a companion matrix is a the diagonal 1's until the *p-1* column(s). In our case this means that the final five columns, i.e., the last lag length, is without any 1's in the diagonal.

Each of the five first rows can be interpreted as the coefficients in the linear regression estimated on the variables. For example, the coefficient value for the second lag of actual house price on user costs, is 0.42, specified in row 3, column 6.

**Table A6.1:** Companion Matrix

| | $p_1^a$ | $r_1$ | $\gamma_1$ | $y_1$ | $h_1$ | $p_2^a$ | $r_2$ | $\gamma_2$ | $y_2$ | $h_2$ | $p_3^a$ | $r_3$ | $\gamma_3$ | $y_3$ | $h_3$ | $p_4^a$ | $r_4$ | $\gamma_4$ | $y_4$ | $h_4$ | $p_5^a$ | $r_5$ | $\gamma_5$ | $y_5$ | $h_5$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $p^a$ | 0.39 | 0.15 | −0.27 | −0.02 | 1.13 | 0.05 | −0.04 | 0.27 | −0.02 | −0.20 | −0.03 | −0.00 | 0.01 | 0.04 | −0.47 | −0.24 | −0.08 | −0.13 | 0.09 | −0.26 | −0.09 | −0.22 | −0.01 | 0.11 | −0.08 |
| $r$ | −0.44 | 0.27 | −0.34 | −0.11 | 3.25 | 0.12 | −0.11 | 0.34 | 0.08 | −2.50 | −0.02 | −0.02 | −0.03 | 0.04 | 1.39 | −0.33 | 0.24 | −0.11 | 0.09 | −2.36 | −0.15 | 0.30 | 0.02 | 0.04 | 0.95 |
| $\gamma$ | 0.46 | 0.18 | 0.82 | −0.08 | 2.46 | 0.42 | 0.05 | −0.06 | 0.14 | −2.37 | −0.04 | 0.10 | 0.10 | 0.09 | 1.32 | −0.41 | −0.42 | −0.24 | −0.01 | −1.49 | 0.19 | 0.01 | 0.00 | −0.01 | 0.57 |
| $y$ | 0.14 | 0.46 | −0.26 | −0.60 | 0.58 | 0.28 | 0.14 | 0.09 | −0.01 | −0.09 | −0.44 | −0.04 | 0.16 | 0.19 | 0.87 | −0.03 | −0.14 | −0.05 | −0.01 | −2.24 | 0.07 | −0.17 | −0.14 | −0.08 | 1.11 |
| $h$ | 0.03 | 0.01 | −0.00 | −0.02 | 0.79 | −0.01 | 0.01 | −0.01 | −0.02 | 0.02 | 0.02 | 0.01 | 0.01 | −0.01 | −0.01 | −0.01 | 0.01 | −0.00 | −0.00 | 0.57 | −0.00 | 0.02 | −0.00 | −0.00 | −0.48 |
| 6 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 20 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 21 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 23 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |

Table A6.1 displays the complete companion matrix with 5x5 rows and 5x5 columns where the five VAR model variables are influencef by past lags.

# A7 Coded Restrictions

Below, in equation (.29) we have plotted the coded restrictions in our VEC model. Here, we implent restrictions on $\alpha$ to be a vector of 1 and 0, corresponding to the speed of convergence in the model. For our $\beta$ restrictions we allow for a -1, 1 relationship, assuming mean-reversion between the fundamental and actual housing price. Lastly we allow for a constant term at 1. Combined this provides our $\mathcal{H}_{1.1}$ as introduced in the analysis section, which can be seen below:

$$\mathcal{H}_{1.1} : \alpha = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ \beta = \begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{.29}$$

# A8 Unit Roots of 1982 - 1993 Sample

Below we have plotted the eigenvalues from our pre-sample VAR model. The pre-sample goes from 1982 - 1993. As we can see, there are eigenvalues with a modulus above 1. These results, ceteris paribus, would suggest that the model is not stable, and therefore that the coefficient matrix cannot be inverted, and used in creating the companion matrix.

**Table A8.1:** Matrix of Eigenvalues of The Pre-Sample

| | | | | |
|---|---|---|---|---|
| 1.1007671 | 1.1007671 | 1.0641429 | 1.0641429 | 1.0400758 |
| 1.0400758 | 1.0348830 | 1.0348830 | 0.9751212 | 0.9751212 |
| 0.9570151 | 0.9570151 | 0.9479193 | 0.9479193 | 0.9423766 |
| 0.9423766 | 0.9324080 | 0.9324080 | 0.9141851 | 0.9141851 |
| 0.8856885 | 0.8856885 | 0.8824554 | 0.8824554 | 0.8316904 |
| 0.8316904 | 0.8299481 | 0.8299481 | 0.7917967 | 0.7917967 |

Table A8.1 displays the eigenvalues generated by the VAR model for the pre-sample. They are sorted from largest modulus to smallest. An eigenvalue $| > 1|$ signifies the model is not stable