Radboud Universiteit

Aiming to investigate the economic factors that are associated with the development of a housing bubble

An analysis of The United States and the Netherlands

Zwitserloot, P.W.A.J.
S4527755
18 September 2020

Contents

Introduction

Het Parool (2020) quoted the national statistics agency CBS, because Amsterdam's housing prices reached for the first time an average of five hundred thousand euro in the last quarter of 2019. Housing prices are increasing over time, in such a way that housing prices in Amsterdam became twice as expensive in comparison to 2013. According to the Rabobank (2019), the reason for this extreme price raising in the real estate market, is through the low interest rates, which are part of the policy of the European Central Bank (ECB). A second reason for the rise in prices is the growing population in the Netherlands. According to the national statistics agency CBS, the population increase in the Netherlands has been the highest since the beginning of the century (12.8 percent in the period of 1995 to 2020.

However, there are also arguments suggesting that the rising housing prices are irrational and is a bubble. Boelhouwer (2017) defines a bubble in the housing market as prices which do not reflect their fundamental values. These raising housing pricing are coursed by individuals who are willing to pay more for houses than the actual price, because they think that prices will grow further in the future. This phenomenon is called speculation and creates bubbles where values of houses cannot be explained by income, wages or demographic developments.

The aim of this thesis is to investigate what variables are associated with increasing housing prices, and might lead to a bubble in the housing market. The research question is: (1) to what extent can we changes in housing prices, and (2) can we predict a bubble in the housing market? We will perform an empirical analysis using the data of the Netherlands and United states.

The basis for this research is the paper of Bourassa, Hoesli and Oikarinen (2019). They compare different techniques of measuring housing price bubbles within six metropolitan housing markets over the world. The literature overview consists of different theories according to the paper of Bourassa (2019), which calculates the bubble in the housing market. These theories are shortly explained, and also important for the following step is the data research itself. So, these theories are providing different variables which are necessary to make the same sort calculations, but then for more recently data and other countries. This research includes the housing market of the Netherlands and the United States.

Literature overview

The housing market is according to the paper of Black, Fraser and Hoesli (2006) an interesting subject of research. Before the crisis of 2007, the housing market did not receive much attention in the literature (Black, Fraser and Hoesli, 2006). Some research papers described the housing price bubbles and made a distinction between regions and countries. For example, Roche (2001), Roehner (1999) described a speculative bubble in a specific city and not in a country.

According to Sanders (2008), the house prices of different cities in the United States appreciated between the 93 percent and the 137 percent between 1996 and 2006. The housing prices were raising until 2006 in the United States. Housing prices were in 2005-2008 Q2 related to delinquent subprime mortgage rates. House prices are continued to fall, and these subprime mortgage rates are still increasing. According to Mian and Sufi (2011), the housing crisis were partly based on the fact that people increased their defaults. Households become during the house price increases able to lend more easy credit for houses (Mian and Sufi, 2009). Also, homeowners extract almost thirty percent of every dollar increase of their homes. The extra available money was used for consumption and not used for long term investments such as stocks, real estate or used for paying off high credit card balances. The financial crisis had as result that house prices felled by almost fifteen percent by the end of 2007 according to Baker (2008). Also, some cities have almost thirty percent of falling house prices. Helbling and Terrones (2003) found that the effect of housing bubble burst takes twice as much time as a stock burst. Helbling and Terrones (2003) argument is one of the four reasons why the housing market is interesting according to Black, Fraser and Hoesli (2006). The effects of wealth are greater for housing than for financial assets (Case et al., 2005; Benjamin et al., 2004). The differences between the household home wealth and the wealth derived from financial assets. People have different motives about bequeathing their stock portfolios or bequeathing their homes. The housing prices and stock prices changes daily, but the emotional impact of houses are different. People have less focus on the short run by houses, instead of stock markets, because this can tracked in newspapers.

Second, according to the papers of Englund et al. (2002) and Flavin and Yamashita (2002) housing is one of the major assets of households. Flavin and Yamashita (2002) described that especially young households have large holdings of real estate relative to their net worth. This led to situation of high risk. Thus, young adults reduce their risks by pay down their mortgage or buy stocks and/or bonds, because the optimal portfolio includes lower ratios of housing to net worth.

Third, the effect on the bussines cycle of a housing bubble burst is almost twice as big in comparison with a stock burst according to Helbling and Terrones (2003). This paper examined the effect of stock and housing bubble bursts and the effect of the economy. They found that a downfall of the stock

market and equity markets are less influencing the output in a country instead of a house price bust. Also, the banking system is more effected by a crash in the housing market instead of the equity market. The capital-to-asset ratio of banks decreases more and therefore the lending capacities decreases. Private consumption in a country decreases immediately after a price busts in the housing market. A crash in the housing market lasted about four years instead of the 2,5 years in the equity market. Also, the recovery time of the economy lasted twice as long in a housing burst instead of a stock burst.

Last, the motive of buying a house is for consumption instead of investment reasons. Thereby, high transaction costs and illiquidity are associated with buying a house. Also, arbitrage possibilities are limited. Although there is correction of the market towards the 'true' value, this 'true' or fundamental value of housing is a prolonged process and inefficient pricing will be perpetuated in uncertain periods.

Calculating the fundamental value of a house contains multiple difficulties, because the theory of demand and supply for housing have to do with a lack of literature according to Hornstein (2009) and Wheelock and Wohar (2009). Also, the fundamental values of houses are dependent of the definition of fundamental. Black et al. (2006) based the calculations of the fundamental value on present value of the disposable income. Other models use a generally constant discount rate such as the paper of Chan et al. (2001). Modeling the discount rate is crucial in order to identify a bubble, even though there is no discrepancy between the fundamental and the actual value. Price bubbles is relatively difficult to apply in practice due computational complexity (Black et al. 2006).

According to the paper of Stiglitz (1990), a bubble in the housing market is where house price levels depart from their 'fundamental' values. There is a lack of literature with regards to calculating housing price fundamentals, but Lind (2009) suggests that a bubble is defined as an extreme price increase, which is followed by a dramatic decrease in prices. There are many different methods to calculate the house price bubble. The paper of Bourassa (2019), makes use of different theories. The first method is based on various ratios according to the paper of Himmelberg, Mayer and Sinai (2005). Ratios which are used in this paper are house prices compared with rents and incomes. The second theory is based on the framework of Abraham and Hendershott (1996), which based their model on housing supply and demand or asset pricing. The third method is based on growth rates in prices (Zhou and Sornette, 2006). An alternative approach to calculate house price bubbles is to calculate the long run relationship between house prices and determinants of the supply and demand for housing. In the next session, these methods will be shortly explained and the next is step is to provide some data. The theories are providing variables which are important for collecting data to conduct data research.

Methodology

Bourassa (2019) proposed three types of methods. The first method is called the present value method. This method calculates the fundamental value of houses. These values could be compared with actual values, in order to calculate if there is a bubble in the market. Second, different ratio measure are used such as price–rent, price–income and imputed–actual rent. These ratios could identify bubbles according to different papers. Last, the regression method is by the Bourrassa et al. (2019), which is useful in order to do regression analysis later in the thesis. The regression is to identify different variables which are influencing house prices. In the following paragraphs, the three types of methods will be discussed which will be used in this thesis to identify bubbles.

1. Present value method

This model is based on the work of Campbell and Shiller (1988a, 1988b) and Black et al. (2006). The present value method tries to estimate the coefficient of the expected future cash flows, which is called the discount factor. This discount factor reflects in the model of Campbell and Shiller (1988a, 1988b), the rate at which future dividends are discounted. The present value model calculates the fundamental value of houses with the key determinants: income, real house prices and interest rates. Also, the use of logarithmic dividend stated as the optimal forecast of the present value of discount factors. These discount factors are useful in the paper of Bourrassa et al. (2019) to summarize the market relevant information and to discount future cashflows.

The paper of Black, Fraser and Hoesli (2006) has been used for the rental growth. The paper of Black et al. (2006) describes an innovative way to compute the fundamental value of housing prices. These values are calculated with the present value of the real disposable income. This model is different than other models, because those are most of the time based on a constant discount rate. These models are based on employment, real income, real construction costs and interest rates of a certain period. The paper of Chan et al. (2001), is modeling the housing market by doing these types of calculations in order to fins a discount factor which could be used to calculate identify a housing bubble. Black et al. (2006) identifying the bubble with an analysis of deviations from the fundamental value of a house. They modelled the bubble component, which relates to the fundamental value and highlight whether a bubble still exist.

The paper of Bourassa et al. (2019) make use of the findings and the methods of the Black et al. (2006) framework. This framework is based on the assumption that the real value of representative agents' residential property (V_t), will be discounted. There will be a constant proportion, named γ which is a portion of the expected future value of real disposable income. The expected future value of

real disposable income is Y_t, is discounted at the real, possibly time-varying, discount rate ρ^*. The ρ^* represents the required rate of return by householders. So, the formula for the real value of household property can be writes as:

$$V_t = \gamma E_t \sum_{i=1}^{\infty} \left(\frac{1}{\prod_{j=1}^{i} (1+\rho_{t+j}^*} \right) Y_{t+i} \quad (1)$$

Black et al. (2006) use an aggregate house-price index, because the formula of (1) holds for the total economy. The constant proportion γ will be replaced by an aggregate house-price index. This is assumed as a constant relationship between the house-price index, P, and market capitalization, V. There has been a constant relationship between the value of all income, Y, and the index which is described above. This leads to a different formula as (1), and the new formula is re-written as:

$$P_t = E_t \sum_{i=1}^{\infty} \left(\frac{1}{\prod_{j=1}^{i} (1+\rho_{t+j}^*} \right) Q_{t+i} \quad (2)$$

Where P_t is equal to $\beta' V_t$ and beta is defined as $\beta'(\gamma)$, and Q_t is defined as βY_t. So, the relationship between income and house prices in the second equation become this: $P_t = \beta'(\gamma) Y_t$.

According to the paper of Meen (1996), the long run ratio of house prices is not constant, because different shocks could influence these house prices. Shocks could influence the relation between income and prices and thereby the cost of capital is also an important factor for the effect of house prices on the long term. So, in this research, the cost of capital will not be constant. However, the cost of capital is unlikely to vary substantially in practice. There has been less variation of the cost of capital over time according to the paper of Black et al. (2006). Vélez-Pareja, Ibragimov, & Tham, (2008) and Arnold & Crack (2004) assume that the costs are constant for simplicity reasons.

The second equation provides a model which forms a basis for the calculation of fundamental house prices. The real fundamental value is modelled as the expected value of future real disposable income which is discounted with the real discount rate (pt). Also, interest rates and real income are being key determinants of real houses. These key determinants will be further explained in the following chapters. The particular solution to $P_t = (P_{t+1} + Q_{t+1})/(1+P_t)$. In order to calculate the real discount rate, the formula become: $1 + p_{t+1} = (P_{t+1} = Q_{t+1})/ P_t$ were P_t is the fundamental value of housing at time t and Q_{t+1} is the real disposable income measured during t + 1.

In order to represent the natural logarithms and their upper-case counterparts the formula could be written as:

$$r_{t+1} = \ln(1 + \exp(q_{t+1} - p_{t+1}) + p_{t+1} - pt \quad (3)$$

In this formula, r is defined as $\ln(1+p)$ and $(q-p)$ is the income-price ratio of the economy. k and μ are both linearization constants based on the income-price ratio. Black et al. (2006) uses the first-order Taylor's approximation, which can be written as:

$$r_{t+1} = -(p_t - q_t) + \mu(p_{t+1} - q_{t+1}) + \Delta q_{t+1} + k \quad (4)$$

Where $k = -\ln\mu - (1-\mu) * (q-p)$

And, $\mu = \frac{1}{(1+\exp(q-p))}$

In this case, $0 < \mu < 1$ and according to the paper of black et al. (2006) is the μ in practice often close to 1. Also, q and p are both 1 in order to ensure stationarity, because variables are transformed. The new formula includes the (log) price income ratio, $p_t - q_t$ and is written as:

$$pq_t = k + \mu pq_{t+1} + \Delta q_{t+1} - r_{t+1} \quad (5)$$

This formula could be repeated substituted for $pq_{t+1}, pq_{t+2}, pq_{t+n}$ on the right-hand side of the equation. Also, $i \to \infty$ and black et al. (2006) assumed that the limit of the last term is 0. So, in this case the model has been linearized, expressed and the conditional expectation of both sides are done by black et al. (2006). The formula has written as:

$$pq_t = \frac{k}{1-\mu} + \sum_{j=o}^{\infty} \mu^{j+1} \Delta q_{t+j+1} - \sum_{j=0}^{\infty} \mu E_t r_{t+j+1} \quad (6)$$

The $E_t r_{t+j+1}$ is the investors required return.

So, the second formula provides a basis formula in order to calculate the fundamental value of house prices. The vector autoregressive (VAR) calculate the expectation of income growth which is based on the work of Campbell and Shiller (1988a, 1988b). The time-varying risk premium is used in order to calculate the expectations on the right side of the sixth formula.

The model of Merton (1973 and 1980) has been used in order to model the time-varying risk premium as a product of the coefficient of relative risk aversion, which is called α. The expected variance of returns is $E_t\sigma_t^2$. The equation for the price-income ratio becomes:

$$pq_t = \frac{k-f}{1-\mu} + \sum_{j=o}^{\infty} \mu^{j+1} E_t \Delta q_{t+j+1} - \alpha \sum_{j=0}^{\infty} \mu^{j+1} E_t \sigma^2_{t+j+1} \quad (7)$$

In this formula, f is the constant real risk-free component of the real required returns. According to Black et al. (2006) the assumption is that the required return is a constant risk-free rate and a time varying risk premium. The time varying risk premium is based on the CAPM work of Merton (1973,1980) which is a product of coefficient of relative risk aversion β, and the expected variance of

housing returns, σ_t^2. The vector autoregressive (VAR) is used to estimate the expected returns, which was introduced by Campbell and Shiller (1988a, 1988b). Black et al. (2006), use the VAR method to forecast the income growth and the housing return variance by this formula:

$$Zt = (Pq_t, \Delta q_t, \sigma_t^2) \quad (8)$$

The ARMA – GARCH model is used by Bourassa (2019) to forecast the housing return variance. The results were very similar between the GARCH model of housing returns and the intertemporal CAPM (black et al. 2006). The empirical VAR is written as: $Z_{t+1} = \boldsymbol{A}z_t + \varepsilon_{t+1}$. In this case the A is the three by three matrix and the ε is the error terms of the vector. Forecasts are realized by multiplying the Z_t by the $J^{th} + 1$ to the power of **A**. The formula become: $Z_{t+1} = \boldsymbol{A}^{j+1}\boldsymbol{z}_t$. The equation which is computed by Black et al. (2006) in order to calculate the fundamental house price and is written as:

$$pq_t^* = \frac{k-f}{1-\mu} + (e_2' - \alpha e_3')A(I - \mu A)^{-1}z_t \quad (9)$$

Where $e_2'A^{i+1}\,z_t = E_t\Delta q_{t+j+1}$ and $e_3'A^{i+1}z_t = E\sigma_{t+j+1}^2$, where e_2' and e_3' are the second and third vector of the vector autoregressive (black et al. 2006). The fundamental value is based on the present value model and the forecasting assumptions. Also, the pq_t^* calculates the fundamental house price after the VAR has been estimated the coefficients and the constants $\mu, k\ and\ r$. So, the fundamental house prices are based on the prediction of income growth which can be written as:

$$P_t^* = pq_t^* + q_t \quad (10)$$

The tenth formula could also be used to calculate how far the actual house is price is from their fundamental value.

---

2. Ratio measures

The paper of Bourassa et al. (2019) makes use of different ratio measures. The measures they are using for measuring bubbles in their paper are price–rent, price–income and imputed–actual rent.

Price-rent ratio is $\frac{Pmt}{Rmt}$ where P_{mt} the real median house price is and the R_{mt} the annual rent is for area m and time t. Thus, this measure the price of a house divided by the rent. The ratio rent-price is calculated by Ambrose, Eichholtz and Lindenthal (2013) for Amsterdam in the period from 1650 to 2005. As result, they found several extended deviations from the average. The price-income ratio is $\frac{Pmt}{Ymt}$, where Y is the real income per capita. According to the paper of Case and Shiller (2003) is the

price income ratio an important measure in order to calculate peaks. This research is conducted in the Unites States and the price income ratio exceed their long- term average by at least 20 percent. level: $R^*_{mt} = P_{mt}E(u_{mt})$, where R^* is annual imputed rent. The last measure is the imputed-actual rents ratio, which is researched by Himmelberg Mayer and Sinai (2005). They compared imputed- actual rent ratio with the other two ratios and found no evidence for bubbles for the imputed ratio. Their results found evidence for bubbles in the period from 1980 to 2004 with the ratio measures price rent and the price income. Imputed rents are defined as the expected user cost per dollar of owner-occupied housing for a typical household multiplied by the real price. E is the expectations operator and u is the annual user cost. These costs differ from area according to Bourassa et al (2015), because this is based on different facts, such as tax, capital gains and tax deductibility and other costs. For example, in the case of Swiss citizens whom are property owner, could deduct taxes. These different types of costs lead to formulas which differ per region. Bourassa et al. (2019) formula for the user cost per dollar of investment in owner occupied housing in different regions. These regions are Finland, Swiss, United stated and the United Kingdom. This research is highly interested in the United States and the Netherlands. So, the formula for the United states according to the paper of Bourassa (2019) is written as: $E(umt)=(1-\tau mt)(imt+\lambda m)+\delta m - E(gmt)$, where τ is the marginal income tax rate for the median household and the capital tax rate. i is the nominal mortgage interest rate, δ is a set of costs not itemized separately (depreciation, property tax, insurance and transaction costs) and g is the nominal house price appreciation. Bourassa et al. (2019) assumes that the cost of equity and debt financing are equal (i). The reason for this assumption is to avoid any assumptions about debt ratios.

In the case of the Netherlands, the Swiss case of Bourrassa et al. (2019) comes close. So, this formula is: $E(u_{mt}) = (1-\tau_{mt})(i_{mt} + \lambda_m + \psi_m) + \delta_m + \tau_{mt}\eta_m - (1 - \tau^{\delta}_{mt})\, E(g_{mt})$

The Swiss property owner could deduct taxes (λ) and maintenance (ψ). In this case, the δ refers to depreciation and annualized transaction costs.

3. Regression methods

The regression attempts to identify house price bubbles according to Bourrassa et al. (2019). According to the theory, the supply and demand of houses are dependent on different variables. For example. Income and interest rates on the demand sides. The supply side is more dependent of construction costs, regulatory constraints (Nitrogen discussion in the Netherlands) and topography. The model in the Bourrassa et al. (2019), calculates the real house prices with different variables which has been found in literature, such as income, population, unemployment rates, real interest rates, real construction costs, rate spreads between 10-year and 3-month government securities and consumer sentiment indexes:

$P_{mt}= (Y^{a}_{mt}, N_{mt}, u_{mt}, \iota_{mt}, b_{mt,} s_{mt}, \omega_{mt})$,

Where p is the natural logarithm of house prices, Y^{a} is the real aggregate income, N is population, u is the unemployment rate, ι_{mt} is the real mortgage interest rate, b is an index of construction costs, s is the term spread and ω is a measure of consumer sentiment.

The second regression method which is used in the paper of Bourrassa et al. (2019) is the EGR approach. This approach focuses on identifying and modeling rates of growth. These rates are considered to be not sustainable. Positive feedback of assets markets can cause price growth to accelerate to unsustainable rates. The formula for this EGR approach is: $p_t= A + B(t_c - t)^m$, where p is the natural logarithm of house prices, t_c is an estimate of the critical point (i.e. peak of the bubble). A, B and m are coefficients. This model could identify a bubble, if P_t increases faster than exponential.

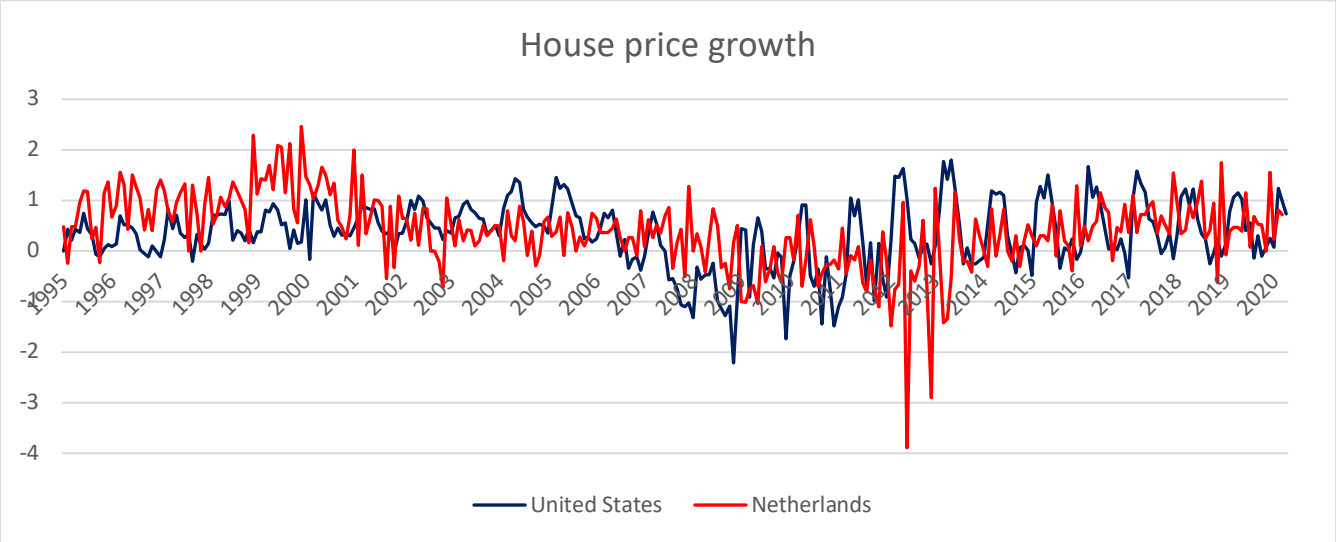
Data and statistical approach

The housing markets which are included in this research are the United States and the Netherlands. The period of 1995 to 2020 will be subject of research. This period reflects different states of the economy. Therefore, the dotcom bubble, financial crisis and a raising economy are included. Bourrassa et al. (2019) did also research to the United States and some other European countries with respect to the housing market.

This thesis has included three types of variables, which are the dependent, independent and the control variable. Bourrassa et al. (2019) uses the house price index as a dependent variable in the model. The house price index which is used in this paper for the Dutch as the United States market indicates the price volatility of the housing prices in a region. The monthly Dutch house price index was derived from the database of the Centraal Bureau Statistiek (CBS). The monthly house price index of the United States was derived from the Federal Housing Finance Agency database.

Dependent Variable

The dependent variable is the House Price Index, for both countries is this monthly data. Also, for both countries this is national level-based data which is in line with the research of black et al. (2006) The data from the Netherlands is based on the Dutch housing prices. The CBS defines this variable as follow: 'the price developments of the current stock of existing owner-occupied homes'. The price index contains every existing owner-occupied home, which is registered in the Netherlands. The start of the index is in 1995 until mid 2020, and 2015 is 100.Graph 1 shows an overview of the growth rates of the housing market. The graph shows that growth rates between the United States and the Netherlands are not really different to each other. According to the article of Bourrassa (2019) the house price data is derived from the U.S. FHFA. This organization notices all transactions in metropolitan area. These transactions are based on single-family homes and those mortgages are securitized by the government enterprises Freddie Mac and Fannie MAE. The provided mortgage is usually for the middle price ranges mortgages, because these organizations are prohibited for the higher segment of houses. Also, low-income homeowner ship is prohibited for these two organizations, which leads to the middle price mortgages. The index is based both for sales as refinancing. Refinancing is depended on the current market, because low mortgage interest could lead to refinancing homes in order to decrease their obligations. For example, in 2001 and the beginning of 2001 consumers start to refinancing mortgage rates. Canner, Dynan & Passmore (2002) suggest that homeowners of the U.S. have taken advantage of the low interests. Lower mortgage rates decreased the monthly obligation, which could possible lead to an increase in saving. Less income is necessary to servicing their mortgage debt. Also, homeowners which refinancing from montages will liquefy some of their equity. In this case, homeowners could possible pay off their former mortgage or use the cash-out for repay other debts, improve their homes or consume other goods. According to the paper of Canner, Dynan & Passmore (2002), refinancing is often after years of apricated house prices.



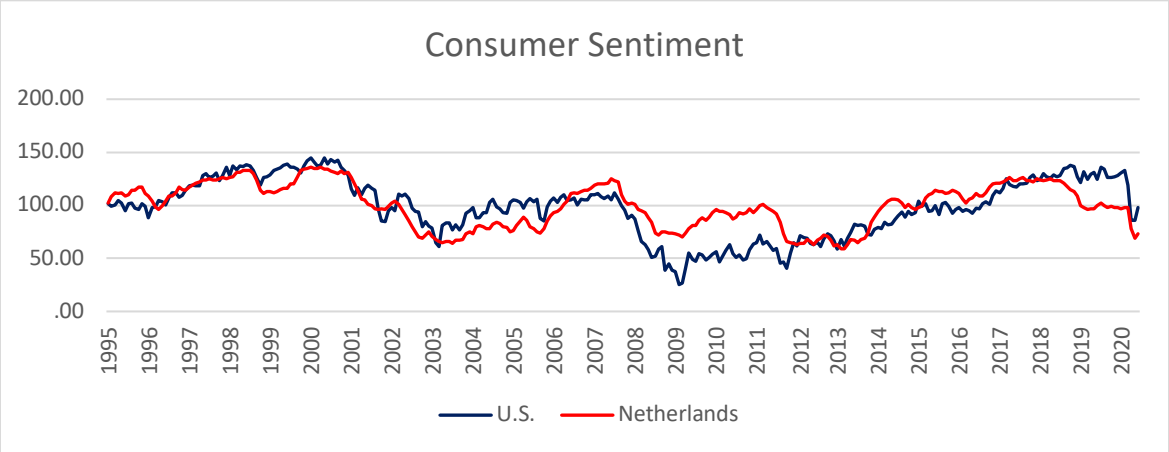
Graph 1: Overview of house price growth in the period of 1995 to 2020 and contains monthly data. The Red line represents the Netherlands, the United States are represented by the Blue line.

Independent and control variables

The independent and control variables that are studied in this thesis are term spread, consumer sentiment, mortgage rate, construction costs, rent, income, number of citizens and unemployment.

The term spread is the difference between the long-term bond yield and the short -term bond yield. According to Bourrassa et al. (2019) the long-term bond yield has been ten years and for the short term three months. So, the term spread has been calculated as the difference between the long-term and the short-term government securities. The U.S. data was found at the database of the department of economic research of the Federal reserve bank of St. Louis (2020). In the case of the Netherlands, the term spread of Germany has been chosen, because more date was available. According to the influential work of Stock and Watson (1989), where they construct a new index of economic leading economic indicators. They found 55 macroeconomic indicators; combine them to predict future economic activity. The national bureau of economic research lists the leading indicators of this selection. One of the leading variables is the spread between the short-term and the long-term U.S. Treasury bond. According to the paper of Estrella and Hardouvelis (1991) and Estrella and Mishki (1997), examining data over the period of 1955 to 1988 and 1973 to 1994 to the predictive power of short- and long-term treasury bonds and conclude that this is a useful predictor for economic growth. Also, this model was useful to calculate the probability of a recession.

The second independent variable that will be used in the analysis is consumer sentiment. Graph 2 reflects the consumer sentiment index in the United States and the Netherlands.



Graph 2: Consumer sentiment index. Data represents the period of 1995 to 2020 and contains monthly data. The Red line represents the Netherlands, the United States are represented by the Blue line. Most data were provided by CBS and Factshet (See appendix for more information).

Consumer sentiment is influenced by economic and financial variables according to the paper of El Alaoui Bouri, & Azoury (2020). Also, Lahiri and Zhao (2016) did some research after the effect of consumer sentiment on the business-cycle of an economy. They suggest that especially in the largest economy, U.S, the consumer sentiment influencing the economy. Van Giesen and Pieters (2019) found empirical evidence that a decrease in economic activities, such as a financial crisis affect the consumer sentiment. This leads to a higher decease in economic activities, which make this a multiplier effect, and even their perceptions of the overall economy and their personal financial conditions are negatively influenced. The effect of consumer sentiment and the housing prices is described by Wilcox (2017), where consumer sentiment is influencing the housing price on the short term and the long term. Consumers decisions about selling or buying a house are based on expectations about the future. The future expectations are influenced by developments in the economy and recent news.

The mortgage rate of the Netherlands was found in the database of the Central bank of the Netherlands. The monthly mortgage rate of the U.S. is from the FRED. According to the paper of Bourrassa (2019), the mortgage interest rate is negatively related to the housing prices. Mortgages rates are in the Netherlands tax deductible. This influences the formula of the Netherlands. Bourrassa (2019) assumes that the costs of financing equity and debt are the same. These costs are both after tax, because alternatives investments would be taxed. Also, mortgage interest rates are deductible from income. The tax system of the Netherlands is quite similar than in the United States, but the property taxes are generally deductible in the U.S. In the United States are maintenance and property deductible.

Construction costs index are related to real costs of new housing which include not only building costs. According to the paper of Capozza, Hendershott, Mack & Mayer (2002), are also restrictions of constructions influencing the costs of new housing. Restrictions such as regulations and policy by the (local) government are influencing the cost of new housing. Also, the stronger restrictions influencing the costs of existing housing. Restrictions have two effects on housing prices, because first the price of new housing will increase and second the ability of builders to respond to demand shocks is reduced (Capozza et al. , 2002) According to Mayer and Sommerville (2000) is that construction is less responsive to price shocks, where more local regulation is.

Also, the unemployment rate, which is found at the CBS (the Netherlands) and FactSet (United States). The unemployment rate includes people between 15 and 75 years old, with institutional citizens excluded. People who have are unemployed and recently have been actively searching for a job are included to the labor force, according to the CBS (2020). The unemployment rate is calculated as number of unemployment divided by the labor force.

Income is the total amount of income in a certain year for one person on average in the United States or the Netherlands. Income is the sum of salaries profits, interest payments, rents and other forms of earnings received in a period (Case & Fair, 2007). The data of the United States is derived from factset. The data from the Netherlands is derived from the central bureau of statistic of the Netherlands.

Also, numbers of rent were found in the database of the Central bureau of Statistics in the Netherlands. Data of the United States was found in the database of the FRED. Rent is an agreement between the property owner and a person for the temporary use of a property. The variable rent is a nominal and a relative value. For example, the CBS has an also data about the yearly rent increase. Some calculations are based on relative numbers, and others are based on nominal values. The calculation of the ratio house price divided by yearly rents are based on nominal numbers instead of growth percentages.

Hypothesis development

Based on the literature, this paragraph shows hypotheses about the effects of independent and control variables on the dependent variable. In the previous paragraph, the variables have already been explained, but this paragraph will elaborate more on the potential effects of the independent variable on the dependent variable. This thesis conducts three types of calculations in order to answer the research question of this thesis. This paragraph is divided in to three parts, because the theory part of this thesis discussed three methods in order to define variables who are associated with house price

increasing and if a housing bubble could be predicted. The three methods which are used in this thesis is a present value method, ratio measures and regression method.

Present value method

The first method of this thesis in order calculate a housing price bubble is the present value method. The present value method calculated the actual and the fundamental value of houses. Also, this calculates how far the actual value of houses are from their fundamental value. Based on the paper of Bourassa et al. (2019), the bubble criterion is when the price of actual houses are twenty percent above the fundamental value. According to the paper of Bourasssa et al. (2019), the expectation is that this happened in the mid 2000s, because this research compares cities in the United States with European cities. In both categories, a bubble was found. First the hypothesis for the Netherlands. So, the hypothesis is as follows:

H0: Actual house prices between 1995 and 2018 in the Netherlands are not twenty percent above the fundamental value
H1: Actual house prices between 1995 and 2018 in the Netherlands are twenty percent above the fundamental value.

The hypothesis for the United States at the present value method is as follows:

H0: Actual house prices between 1995 and 2018 in the United States are not twenty percent above the fundamental value
H1: Actual house prices between 1995 and 2018 in the United States are twenty percent above the fundamental value.

The effect of different variables on the house price index could be positive or negative. According to the paper of Bourrassa et al. (2019), a model has been created to predict the effect house price index across time. The model is as follows:

$P_{mt} = (Y^{a}_{mt}, N_{mt}, u_{mt}, \iota_{mt}, b_{mt,} s_{mt}, \omega_{mt})$

This model in calculates the real house prices with different variables which has been found in literature, such as income, population, unemployment rates, real interest rates, real construction costs, rate spreads between 10-year and 3-month government securities and consumer sentiment indexes. The regression attempts to identify house price bubbles according to Bourrassa et al. (2019).

Ratio measures

The second method of this thesis is the method of ratios. The ratios which are used in this thesis are the price – rent, price-income and the imputed-actual rent. First of all, a hypothesis of the price – rent ratio should be developed. The price – rent calculates the ratio between the price of a house and the yearly rent of property. According to the paper of Case and Shiller (2003) is a exceed of 20 percent with regards to the long- term average enough to calculate a bubble. According to the paper of Bourassa (2019), was there a bubble in the mid 2000s. So, the hypothesis is as follows for the Netherlands:

H0: The ratio price-rent is between 1995 and 2018 in the Netherlands not twenty percent above the long-term average.
H1: The ratio price-rent is between 1995 and 2018 in the Netherlands twenty percent above the long-term average.

The hypothesis for the United States at the present value method is as follows:

H0: The ratio price-rent is between 1995 and 2018 in the United States not twenty percent above the long-term average.
H1: The ratio price-rent is between 1995 and 2018 in the United States twenty percent above the long-term average.

The second ratio is the price- income ratio. This ratio is the house prices divided by the yearly income in a country. According to the paper of Case and Shiller (2003) is this ratio an important measure in order to calculate bubble if the long-term average exceeds by 20 percent. According to the paper of Bourassa et al. (2019), there was a bubble in Europe and in the United states. So, expected is a bubble in the mid 2000s. The hypothesis is as follows:

H0: The ratio price- income is between 1995 and 2018 in the Netherlands not twenty percent above the long-term average.
H1: The price- income is between 1995 and 2018 in the Netherlands twenty percent above the long-term average.

The hypothesis for the United States is as follows:

H0: The ratio price- income is between 1995 and 2018 in the United States not twenty percent above the long-term average.

H1: The ratio price- income is between 1995 and 2018 in the United States twenty percent above the long-term average.

Regressioin analysis

Last of all, the expected effect of the dependent variables is explained with different literature to develop different hypotheses. After the development of these hypotheses, the statistical analysis will be performed with STATA. According to Ahrens (2002) are term spreads a reliable predictor of the business cycle peaks and throughs. This model could possibly forecast the business cycle. This model is based on the two-state Markov process. According to the paper of Wheelock and Wohar (2009), most literature describes the relationship between the yield curve and the economic activity, and that the time spread forecasts growth or recessions. The problem is that there is less literature available for the reasons why this time spread is forecasting the economic activity. The forecasts are useful for a time horizon of six to twelve months. According to the paper of Dotsey (1998) movement at the spread changes the real GDP. Also, the term and the GDP are positively correlated with each other. Thus, the spread indicates whether if the GDP will grow or will decrease over time. The business cycle is also predictable by the time spread and also positively related to the GDP of a country. So, in this case the hypothesis will be:
H0: Term spread has no effect on the house price index
H1: Term spread hasan effect on the house price index

The mortgage rate has a negative relationship according to Abelson et al. (2005). If mortgages become more expensive, people are less likely to buy a house and so the house price index will decrease. In the case of mortgage rates, Abelson et al. (2005), there appears a strong negative relationship; a 1 per cent rise/fall in the mortgage rate will lead to a fall/rise in real house prices of 5.4 per cent on average.

H0: Mortgage rate has no effect on the house price index
H1: Mortgage rate has a negative effect on the house price index

Consumer sentiment according to the paper of Johnson (2010) has a positive relationship with the house price index. So, if the consumer sentiment increases, the prices of houses are followed by an increase. Sentiment can be defined as the attitude of consumers or the consumer optimism or in lower levels the pessimism. Pessimistic consumers are less likely to make purchases related to consumption and simply consume, because they tend to save. Dua (2008) also investigated the effect of the consumer sentiment on the house market. Dua (2008) demonstrated that consumers based their sentiment on different variables, like expected interest rates, wealth, financial status and current prices of homes. The Home Buying Index is created by Dua (2008) and indicates the consumer sentiment about the housing market. The formula as follows:

$$Home\ \ Buying\ Index = Good + Uncertain\left(\frac{Good}{Good+Bad}\right) \quad (10)$$

This index calculates a value between zero and hundred, which indicates consumer feelings about the buyer conditions. This index is positively related to the house price index, because an increase in the Home Buying Index meant that the conditions for buying a house is increased (Dua, 2008). So, in order to develop a hypothesis based on the works of Johnson (2010) and Dua (2008) is as follows:

H0: Consumer sentiment index has no effect on the house price index
H1: Consumer sentiment index have a positive effect on the house price index

According to the paper of Case and Shiller (1990), construction costs are positively correlated with the house price index. House prices will rise after construction raise, because houses become more expensive to build. Also, Hendershott and Abraham (1993; 1994) found that construction costs and income have a positive effect on the house price. Capozza et al. (2002), are using constructions costs as an indicator for the short-run responsiveness of supply to demand shocks. The constructions costs are based on building costs and the degree of restrictions, which are positive related to the house price index. The hypothesis is as follows:

H0: Construction costs have no effect on the house price index
H1: Construction costs have a positive effect on the house price index

Englund and Ioannides (1997) suggest that the number of citizens is positively correlated to the house price index. Especially, when the number of citizen's are lagged according to the house price index. According to the paper of Abelson, Joyeux, Milunovich and Chung (2005) unemployment rates negatively influence the house price index. If the unemployment rate increases, the income will decrease, and people are less likely to purchase a house. Also, people are less likely to receive a mortgage from a bank in the case of uncertainty in the market. The demand of houses is negatively correlated with the unemployment rate. So, there is a negative relationship between the unemployment rate and the house price index.

According to the paper of Jacobsen & Naug (2005), income has a positive relationship with the house price index. Income is one of the most important explanatory variables for house prices. The amount of credit for a mortgage is highly based on future income and housing costs. The future income is dependent on expectations about the future. Housing costs are relevant because housing is durable, and households are debt-financing a substantial portion of the purchase price (Jacobson & Naug, 2005). So, if the income becomes higher, the prices of houses will also increase.

Data Analysis

The collected data is panel data, because it consists of a multidimensional dataset. A requirement for panel data is that there are multiple times of measuring. This dataset includes most of the monthly data of a period from 1995 until the middle of 2020. Another requirement of panel data is that the same economic entity has been measured in different time frames. In this paper, the economic entity is the house price, which is measured over time. Two models of panel data are the most common. These models are the fixed effect model and the random effect model. In order to test which model fits the data, the Hausman test will be done. First of all, there should be a zero hypothesis and an alternative hypothesis developed. The requirement for the p-value to be significant, is smaller than 0.05 ($P< 0.05$).

H_0= The random effect model is preferred
H_1= The fixed model is preferred

The P-value is lower than 0.05 if H_0 become rejected and H_1 assumed. If the random effect model fits the data, there should be a test for the variance. The variance test is called Breusch and Pagan test, which calculates the variance in the model. In case if the variance is zero, there is no random effect model. So, the hypothesis is as follows:
H_0= the variance = 0
H_1= the variance ≠ 0

The next step is to control for multicollinearity, because it could be possible that different variables have the same effect on the dependent variable, the house price index. The multicollinearity is measured in two different ways. The first method is the collinearity matrix, because this matrix represents the correlation between two variables. Determining the criteria, whether there is too much correlation between the variables, the Rule of Thumb is needed (Hinkle, 2003). The Rule of Thumb table can be found in the Appendix 2. The correlation between variables is not allowed to be higher than 0.9, because that would resemble multicollinearity. The second possible test for multicollinearity is the Variance inflation factor. According to O'brien (2007) is the VIF the perfect test for multicollinearity in a database. Hair, Anderson Tatham & Black (1995) argue that a database with VIF score lower than ten, don't contain multicollinearity. These tests are both conducted in this thesis.

The formula that is used in the regression analysis in STATA is as follows: $P_{mt}= ß_1 INCOME^a_{mt}+ ß_2 N_{mt} + ß_3 UNEM_{mt} ß_4 R_{mt}, ß_5 CONSTRUCT_{mt,} ß_6 TS_{mt}, ß_7 CS_{mt}$ (11)

Meaning of the variables:
$P_{mt}= (Y^a_{mt}, N_{mt}, u_{mt}, ı_{mt}, b_{mt,} s_{mt}, ω_{mt})$,

P_{mt}= House price index = HPI

Y^{a}_{mt},= income = INCOME

N_{mt}, = Number of Citiziens = N

u_{mt},= Unemployment rate = UNEM

ι_{mt}, = Mortgage rate = R

$b_{mt,}$ = construction costs = CONSTRUCT

s_{mt}= term spread = TS

ω_{mt}= Consumer sentiment = CS

Testing hypothesis for the present value method and the ratio method, another data analysis has to be done. The data contains only yearly data for both methods. First the present value method should be calculated according the methodology part in this paper. This calculates the fundamental value of the house prices in the Netherlands and in the United States. According to the paper of Bourrassa et al. (2019) and Black et al. (2006), the deviations of the actual house price should be calculated. So, in order to calculate this, the difference between the fundamental and the actual value of the house prices should be calculated. Testing the hypothesis with the criteria of Bourassa et al. (2019) which is that twenty percent difference between the fundamental and the actual value is a housing bubble. Second the ratio method which contains the price- income ratio and the price- rent ratio. The calculations are mentioned in the methodology part of this paper. Also, the long-term average is calculated, because testing the hypothesis according the criteria of Case and Shiller (2003), a possible is bubble identified if the actual ratio exceeds the long-term average by twenty percent. The results are discussed and compared with each other.

Results

Summery statistics

This paragraph represents a summary of the statistics that are used in the analysis. Table 1 shows an overview of the data, both of the United states and the Netherlands. The variables which are included are the house price index, income, unemployment, number of citizens, mortgage rate, consumer sentiment, construction costs and term spread.

The house price index has a minimum from hundred, because 1995 has been chosen as basis year. The maximum house price index is 272.23, which is in the United States in 2020. This is a yearly average increase of almost 6.8% for the United States. For the Netherlands, this number is lower and is almost 4.8% per year on average. The second variable is income, which is on average also growing over time. The lowest total income in one year is the 6291.4, which is in 1995 in the United States. The central bureau of statistic in the Netherlands only registered from 2000. The maximum total income is

in the Netherlands in 2018, which is the latest year of registration for both countries. The third variable is unemployment, which had a maximum of ten percent. This percent is in the end of 2009 in the United States, in the middle of the financial crisis. According to the paper Weidner and Williams (2011), the peak of 2009 is related to the reducing incentive of the unemployment to seek and accept fewer desirable jobs. Also, there became degree of mismatch between job seekers and potential employers. Es especially workers in the construction, finance and real estate sector were not easily transferable to growing sectors. According to the paper of Barchnichon et al. (2010), growing sectors as education and health care, are less available for people without education in this sector. The last reason for the rising unemployment numbers in 2009, are come from the homeowners. After the housing bust, millions of homeowners have underwater mortgages. As a result of this situation, homeowners are locked in their homes and make it more difficult for them to move to higher growth areas.

The number of citizens is in both countries a growing number over time. The minimum number of citizens is in the Netherlands in 1995, and the maximum number of citizens is in the United States in 2020. The mortgage rate (R) has fewer observations, because the Dutch central bank doesn't have monthly data in the beginning of the research period. The minimum mortgage rate is 2.9 and is the last number of the dataset of the Netherlands. In comparison with the United States, which have the highest mortgage rate in 1995, have a mortgage rate of 3.72 in 2020. According to Harris and Meier (2015), this difference is for structural reasons and not based of levels of financial distress. The variable consumer sentiment has six hundred observations in total, because this is monthly data. The mean of this data is 98.96, which is almost the start of the index with basis year 1995. The lowest point of sentiment was in the beginning of 2009 in the United States. As mentioned before this was in time of the financial crisis. In comparison with the Netherlands, the sentiment was higher. The sentiment of the Netherlands was in the beginning of 2009, 73. After this lowest point in time, the sentiment was low until the middle of 2013. In the Netherlands, the sentiment is also growing in 2013, but in the end of 2013 instead of the middle. The highest peak is in the beginning of 2000 in the United States, which is before the dotcom bubble burst. The next variable is the construction index, which have as reference 1995. Both countries have lower construction costs during the financial crisis, but after the crisis, these costs are rising. So, this makes an index of 197.55 possible in 2020 for the United States. The Netherlands have in 2020 an index of 155.76 at the end of the dataset. The last variable is the term spread and has a negative number as minimum. The negative term spread is in the end of 2008 in the Netherlands. The United States have positive term spread in the end of 2008. Estrella and Hardouvelis (1991) found that term spread significantly outperforms other financial and macroeconomic variables in order to calculate U.S. recessions. Also, Bernard and Gerlach (1998) found that the term spread also forecasts recessions up to two years in the Netherlands. Wheelock and Wohar (2009) found that the term spread forecast the output differences or recessions, but not the

reason for this. According to Benati and Goodhart (2008), there are forecasts possibilities with the term spread, but there is a lack of theory behind these findings.

Variable	Obs	Mean	Std. Dev.	Min	Max
logHPIL	600	5.161818	.2380356	4.60517	5.606647
INCOME	45	17206.66	7022.828	6291.4	29500
UNEM	508	5.489173	1.53086	2.9	10
N	600	1.58e+08	1.43e+08	1.54e+07	3.29e+08
R	504	5.163909	1.376053	2.9	9.15
CS	600	98.96183	24.28192	25.3	144.7
CONSTRUCT	600	141.5353	24.7588	99.27204	197.5514
TS	600	1.36678	1.033666	-1.2331	3.68

Table 1: Summary of statistics. Respectively HPI (house price index), Income (income), UNEM (unemployment rate), N (number of citizens), R (mortgage rate), CS (consumer sentiment), TS (term spread). Income was yearly data. The remaining variables consist of monthly data over the period of 1995 until 2020.

Hausman test

In order to check which model needs to be used, a random effect model or a fixed effect model, the Hausmann test needs to be done, according to Mutl and Pfaffermayr (2011). In the previous chapter, the hypotheses were already explained. It was expected that a fixed effect model needs to be used, because for several variables more than six hundred observations were included. As a result, the changes are smaller per period, because monthly data has been used. Also, some variables are changing slowly over time in this time period of 25 year. For example, construction is also dependent on policy. According to Herring (2007), introducing and new regulation takes a lot of time.

The results of the Hausman test can be found in table 2 in the appendix. The hausman test showed that the fixed model will be used.

Breusch and Pagan Test

In the last paragraph the Hausman test showed that a fixed model is needed to analyze the data. In this paragraph, the Breusch and Pagan test is done to evaluate whether an OLS regression fits the data better than a panel data approach. (Breusch and Pagan, 1979). Table 3 in the appendix shows an overview of the Breusch and Pagan test. The results indicate that there is homoscedastic in the data,

and this means that the variance of the residuals is independent of the independent variable. This test showed that an OLS regression is needed to analyze the data.

Multicollinearity

Multicollinearity arises if two or more variable interfere with each other. This could lead to results which are deceived. If multicollinearity plays a role in the data, one variable could be seen as sufficient for the regression. Testing multicollinearity is possible in two ways. First of all, the correlation matrix which can be found in the appendix table 4. This table is analyzed with the Rule of Thumb (Hinkle, 2003, see Appendix 2). This shows that there is high collinearity with the number of citizens (N) and income. Also, there is a high correlation between the logged house price index, mortgage rate (R) and construction costs (CONSTRUCT).

The second test for multicollinearity is the variance inflation factor (VIF). According to the paper of O'brien (2007), the VIF test calculates and shows a representation of multicollinearity. The benchmark which is used for this test is also the Rule of Thumb according to Hair et al. (1995). In this case, a value smaller than ten, shows that there is no multicollinearity in the model. Table 5 in the appendix shows that the VIF is higher than 10, which indicates multicollinearity in income and number of citizens.

According to the paper of Neter et al. (1989). Three problems could arise from multicollinearity. First of all, the variance of the coefficient can become unreasonably high. Besides, the opposite outcome can occur in the analysis. Also, the parameters could not show significant effects. There are two possible solutions for this problem. The first option is to drop the variables which are multicollinear. This option would lead to model without income and number of citizens. According to the paper of Bourrassa (2019) income and the number of citizens is part of the formula. So, the research will be influenced by dropping variables.

The second option is that these two variables had to be centered around the mean. The calculation for the number of citizens is as follows:

$$(N - mean)^2 \quad (12)$$

For income is the formula as follows:

$$(INCOME - mean)^2 \quad (13)$$

According to the paper of Aiken and West (1991), Cohen and Cohen (1983), Jaccard, Wan and Turrisi (1990) and Smith and Sasaki (1979) mean centering can be used in order to reduce multicollinearity. After centering the variable INCOME at CINCOME and N in CN, the VIF test doesn't suggest any

multicollinearity. Table 6 show the new results of multicollinearity, which are less than the VIF test before. Also, according to the rule of Thumb, multicollinearity is solved in this model.

Variable	VIF	1/VIF
R	**4.85**	**0.206015**
CINCOME	**4.55**	**0.219918**
UNEM	**2.72**	**0.367244**
CS	**2.53**	**0.394624**
CONSTRUCT	**2.36**	**0.422900**
CN	**1.83**	**0.547140**
TS	**1.71**	**0.583616**
Mean VIF	**2.94**	

Table 6: Results of multicollineairity tests after centering the means.

Regression analysis

The Hausman test concluded that the data could be analysed with a fixed effect model. However, the Breusch and Pagan Test showed that a classical OLS regression fits the data better. Afterwards, the multicollinearity test was positive. After centering around the mean, the VIF was lower than ten, which indicates the multicollinearity was solved. In this paragraph, a regression analysis of the data of the paper is performed and explained.

The regression analysis is needed to identify the house price bubble, according to the paper of Bourassa et al. (2019). This model contains different variable, which divided in to three types of variables (dependent, independent and control). The model is as follows:

$$P_{mt} = (Y^{a}_{mt}, N_{mt}, u_{mt}, \iota_{mt}, b_{mt}, s_{mt}, \omega_{mt}) \quad (14)$$

This thesis describes the logarithmic function of the house price index. According to the paper of Bourassa et al. (2019), p is the natural logarithm of house prices, Y^{a} is the real aggregate income, N is population, u is the unemployment rate, ι_{mt} is the real mortgage interest rate, b is an index of construction costs, s is the term spread and ω is a measure of consumer sentiment.

The first variable is unemployment, which was expected according to the paper of Abelson et al. (2005) to influence the house price negatively. People would be less likely to buy a house if the unemployment rate increase, because more uncertainty is in the market. So, the demand for houses decrease, which influence the house price negatively. As expected according to Abelson et al (2005), the unemployment rates have a negative influence on the house price index. The negative correlation is at a five-percentage significant and has a negative coefficient of 0.035. So, the unemployment rate

influences the house price negatively, which means that if there are more people unemployment, the house price will decrease in a country.

The second variable is the R, which is standing for the mortgage rate of the country. This statistic is found in the Central bank of the Netherlands and the data from the United States is arrived from the Federal Reserve Economic Data (FRED). Bourassa (2019) found that the mortgage rate was negatively related to the house prices. Also, Abelson et al. (2005) concluded that mortgage rates have a negative relationship with house prices. People are less likely to receive a mortgage if mortgages become more expensive. The demand for houses will decrease and this is followed by a drop in the price. The results of this thesis are in line with the expectations based on the articles from Abelson et al. (2005) and Bourassa (2019). However, the results are not statistically significant at 0.05, but are below 0.1. Therefore, there could be trend that the mortgage rate has a negative relationship with the house price index

The third variable is CS which is consumer sentiment in a country. According to the paper of el Alaoui et al. (2012) this variable is influenced by economic and financial variables. Economic activities are influenced by consumer sentiment and affecting the economy (Lahiri and Zhao, 2016). Also, consumer sentiment is affected through a vicious cycle, because if consumer sentiment decreases, economic activities also decrease. This influence the consumer sentiment, which creates a multiplier effect according to Van Giesen and Pieters (2019). Consumers based their short- and long-term consumptions on future expectations (Wilcox, 2017). Expectations are based on developments in the economy and recent news. According to the paper of Johnson (2010), a positive relation between consumer sentiment and the house price was expected. Dua (2008), found that consumers based their sentiment on different variables, like current prices of home, expected interest rates and financial status. According to the results in table 7, the consumer sentiment has a negative relation with the house price index. The coefficient is minus 0.0022 and is significant by a five-percentage significance level. A negative relationship between consumer sentiment and house prices indicates that if consumer sentiment increase, the houses prices will decrease. Table 7 shows also results with lags that indicates a negative relation between consumer sentiment and the house prices.

The fourth variable is CONSTRUCT, which means construction costs. The constructions costs contain real costs of new housing, but also according to the paper of Capozza et al. (2002), restrictions of new buildings. Regulations of the local government are influence the price of new housing. According to Case and Shiller (1990), are construction costs positive related to the house price index, because higher construction costs will lead to higher prices of new build homes. Also, Hendershott and Abraham (1993; 1994), found that construction costs influencing the price positively. The results in

table 7 verify the expectations. The coefficient of CONSTRUCT is positive and is significant at a level of one percent. Thus, if the construction costs increase, the house price will increase.

The fifth variable is TS, which is the term spread. The term spread is calculated as the difference between the long-term (ten years) and the short-term (three months) bond yield (Bourrassa et al. (2019). Stock and Watson (1989) construct a new index of economic leading indicators, which could possibly predict future economic activity. This index includes the term spread and according to the papers of Estrella and Hardouvelis (1991) and Estrella and Mishki (1997), isa predictor of economic growth. Also, this model was useful to calculate the probability of a recession. According to the paper of Dotsey (1998), the spread is positive related to the house price index. The results are shown in table 7 , where the term has no significant relationship between the term spread and the house price index.

The sixth variable is CINCOME, which is the centered income variable. The variable income was centered around the mean in order to solve multicollinearity. According to the paper of Aiken and West (1991), Cohen and Cohen (1983), Jaccard, Wan and Turrisi (1990) and Smith and Sasaki (1979) is mean centering important for reducing multicollinearity in the dataset. According to the paper of Jacobsen et al. (2005), the effect of income is positive related to the house price index. The variable is positively related to the house price index, but the coefficient is really small with a significance level of five percent. So, if the income in a country is rising, the house prices will also raise.

The last variable is CN, respectively number of citizens. This variable was also centered around the mean in order to solve multicollinearity. The number of citizens is for both countries a growing over time. According to Englund and Ioannides (1997) the number of citizens is positively related to the house price index. Table 7 shows that the number of citizens has a negative relationship with the house price index. The coefficient is really small, but negative and was not statistically significant. Therefore, it cannot be stated that the number of citizens has a positive or negative effect on the housing prices.

The results show that the mortgage rate R and the number of citizens does not have a significant effect on the house price index. The number of citizens is not significant, according to Houle & Berger (2014), citizens buy their first home between the age of 25-34. So, the number of citizens affecting the house price probably with lags. According to the paper of Tsatsaronis & Zhu (2004), the Netherlands and the United States are countries where citizens with a mortgage has a fixed contract with regards to mortgage rates. This research compared the Netherlands and the United States with different countries, such as the United Kingdom, Norway and Australia. The short-term mortgage rates changes influence the house price more than 0.8 percent more than the United States and the Netherlands, because of the long-term fixed contracts.

The results showed that unemployment, consumer sentiment and income have a relationship with the house price index with a significant level below five percent. The variable construction costs have a positive significant relationship below one percent. The variable term spread is not significant and have no relationship with the house price index. The adjusted R^2 is generally high, but this comparable with the model of Bourassa et al. (2019). Also, Bourassa et al. (2019) found no explanation for this.

The robustness check is of great importance to check whether the outcomes of the regression analysis are reliable. The robustness check is conducted with a different composition of the data. The results are reported in table 8. The robustness check doesn't change the coefficients. The dependent variable is the logged house price. The unemployment rate was in the regression model significant at a level of five percent, but in the robustness check, the significant level is below one percent. The mortgage rate was in the regression model significant at a level of ten percent, but with the robustness check, the new level of significance is five percent. The variables which stay the same are the consumer confidence, construction costs, term spread, income and number of citizens. The robustness check in this analysis is to check the robustness of the conclusions.

VARIABLES	logHPIL	logHPIL
UNEM	-0.0348**	-0.0348***
	(0.0143)	(0.0127)
R	-0.0372*	-0.0372**
	(0.0204)	(0.0175)
CS	-0.00223**	0.00223**
	(0.000861)	(0.000825)
CONSTRUCT	0.00604***	0.00604***
	(0.000903)	(0.00107
TS	0.00699	0.00699
	(0.0159)	(0.0137)
CINCOME	2.56e-10**	2.56e-10**
	(1.06e-10)	(9.4e-11)
CN	-0*	-0*
	(0)	(0)
Constant	4.824***	4.824***
	(0.290)	(0.288)
R-squared	0.893	0.983

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Regression analysis. Respectively HPI (house price index), Income (income), UNEM (unemployment rate), N (number of citizens), R (mortgage rate), CS (consumer sentiment), TS (term spread). The second column is the robustness check.

Results Present value analysis

This paragraph represents a summary of the statistics that are used with the present value and the ratio method. Table 7 shows an overview of the data for the United States and table 8 for the Netherlands. Differences between the other summery statistics is that these are all yearly data, instead of monthly data, because limits of availability of data. First are the summery statistics of the United States and the Netherlands are discussed and compared with each other in this paragraph.

The logged house price ratio has a minimum of minus 0.11, which was in the crisis (2008), when house prices decreased dramatically. The maximum of the logged house price is 0.077, which were in the years before the crisis of 2007. Comparing this with the Netherlands, the minimum value is minus 0.04 and the maximum is 0.074, which are both smaller than the United States. The second variable is the logged rents in the different countries. The statistics of the United States and the Netherlands are comparable. A small difference between the values is the mean, which is United States 0.013 and in the Netherlands 0.12. The third variable is the logged price rent, which in the United States less volatile than in the Netherlands. The Netherlands has a minimum of 1.38 and a maximum 1.622 and the United States the numbers are between 1.25 and 1.37. Also, the mean and the standard deviation is lower in the United States. The fourth variable is the price income. In the United States, income increased more than house prices in contrast to the Netherlands were house prices increased more than income. Also, the factor of house prices divided by income is in the United States bigger than in the Netherlands. The mean in the Netherlands is 8.86, instead of the 19.54 in the United States. This difference is because the income in the Netherlands is in 1995 almost three times more than in the United States. In 2018, the average income in the Netherlands is only 1.5 more than in the United States. The fifth variable is the house price- rent ratio, which reached their maximum for the Netherlands before the crisis in 2007. The maximum value in the United States was growing before 2018 and reached his maximum by the last measurement. The mean of the Netherlands is higher than the mean of the United States, because rents are in the Netherland substantially lower. According to Haffner & Bouwmeester (2010), the Netherlands have a strong social rental sector, were large section of the population living. The effect of this strong social rental sector is that the average rent is lower than other countries. The variable logged construction is for both countries in the crisis of 2007 negative and the maximum value is for both in 1996. The mean for the United States is higher than of the Netherlands. Also, the standard deviation is higher for the United States. The variable logged income has a mean of 0.18 in the United States and 0.12 in the Netherlands. According to Dynan & Sichel (2012), income in the United States was in the late 2000s more volatile, because of the economic activities in this period. On average income is for both countries growing over time. The crisis in 2007 caused for the only negative growth numbers of income. Some of the variables are also mentioned in the previous chapter. The difference is only the number of observations, because the table in the earlier chapter are monthly data and this table is yearly data.

The next variable is consumer sentiment has higher mean in the Netherlands than in the United States. The crisis in 2007 caused a drop in the consumer sentiment for both countries as mentioned before. Also, unemployment rate obtained his highest number in the middle of 2009. The peak of 2009 is related to the reducing incentive of the unemployment to seek and accept fewer desirable jobs (Weidner and Williams, 2011). Also, according to the paper of Barchnichon et al. (2010), are people locked in their homes, because of their underwater mortgages. As result, people didn't move to higher growth areas. Number of citizens are for both countries a growing number over time. The unemployment rate for the United States are at the highest level in the middle of the crisis in 2007. In the Netherlands was the highest rate after of the crisis in 2013. The mortgage rate is more flexible in the United States, because this is more fluctuating over time and has a standard deviation of 0.15. The Netherlands has a less volatile mortgage rate, because they have a standard deviation of 0.005. This difference is also seen at the last variable: the term spread. The term spread is also more volatile in the United States. The minimum and the maximum are more differ from each other in the United States than by the Netherlands. Also, the standard deviation is higher by the United States.

Variable	Obs	Mean	Std. Dev.	Min	Max
loghousepr~e	24	.0257592	.0425041	-.1114835	.0774769
logrent	24	.0130582	.0040921	.0020443	.0198466
logpricerent	15	1.313269	.0379311	1.251108	1.370821
priceincome	24	19.53716	1.423242	16.87577	22.31983
houserent	15	20.6451	1.807375	17.82821	23.48667
logconstruct	24	.0113286	.0198536	-.0421624	.0433749
logincome	24	.018572	.0094095	-.0140208	.0314904
CS	24	96.05417	26.5575	37.4	144.7
n	24	3.00e+08	1.87e+07	2.65e+08	3.28e+08
U	24	5.804167	1.636135	4	9.8
R	24	.0576083	.0154408	.0341	.09
TS	24	1.671679	1.158477	-.35	3.671579

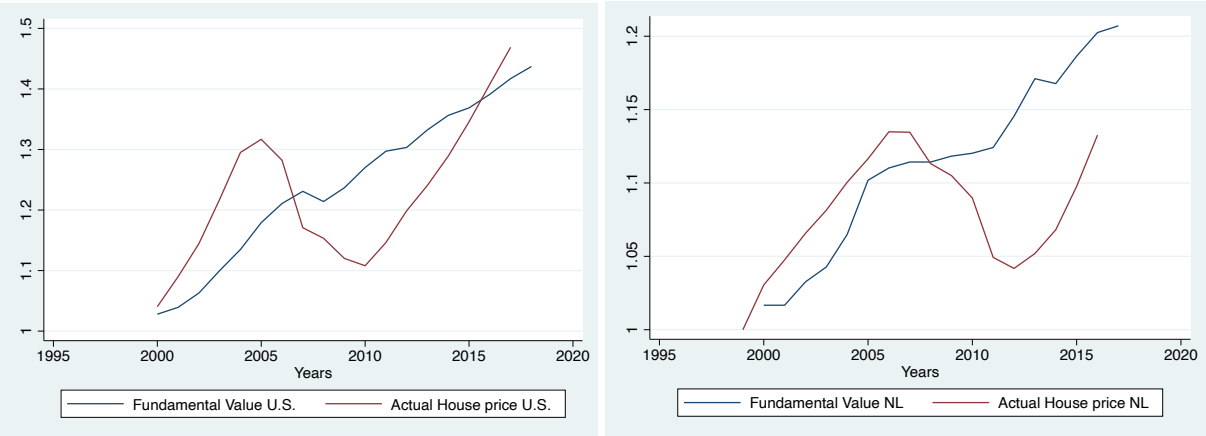
Table 7 Summery statistics of the United States.The variables which are included in the table are log House price, log rent, log price rent, price income, house price rent, log construct, log income, consumer sentiment, number of citizens, unemployment rate, mortgage rate, term spread.

Variable	Obs	Mean	Std. Dev.	Min	Max
loghousepr~e	23	.0190077	.0267008	-.0404848	.073507
logrent	24	.0120456	.0040255	.006038	.0199467
logpricerent	23	1.538851	.0661058	1.381862	1.622341
priceincome	24	8.859409	2.176593	5.700389	17.21645
houserent	23	34.94991	4.977197	24.09142	41.91225
logconstruct	24	.0053342	.0046821	-.0050763	.0194824
logincome	20	.011605	.0097251	-.003037	.0340798
CS	24	100.5583	21.35871	59	135
n	24	1.64e+07	512977	1.55e+07	1.72e+07
U	16	5.39375	1.162451	3.7	7.6
R	16	.0466937	.0048178	.0358	.0565
TS	24	1.209996	.8875138	-.4515	2.66

Table 8 Summery Statistics of the Netherlands

Results present value method

This paragraph is answering the hypothesis about the present value method. The hypothesis is the same for the United States and for the Netherlands. The fundamental price is based on the model of Campbell and Shiller (1988a, 1988b). The fundamental price is calculated with the log price-rent ratio, risk free rate, income and house prices. The actual house price is based on the data which is found. First of all, graph 3 shows the fundamental value and the actual value. A bubble in the house market is according to the paper of Stiglitz (1990), where house price levels depart from their 'fundamental' values. The graph indicates if there is a bubble in the data. According to the paper of Bourrassa et al (2019) a difference of twenty percent is enough in order to identify a housing bubble.

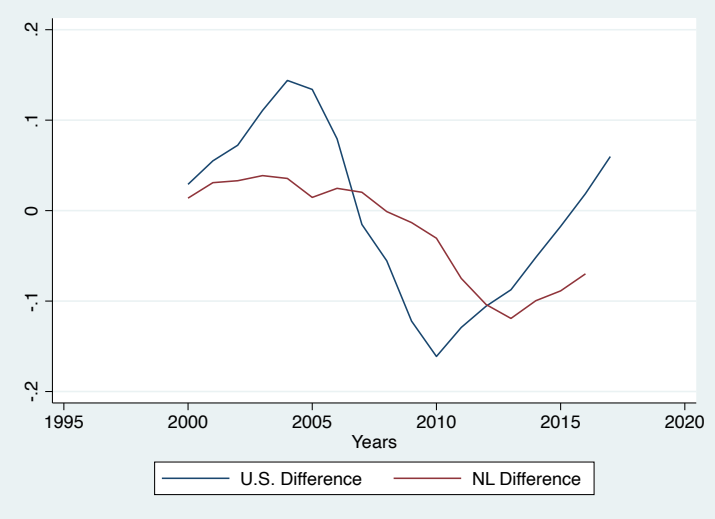


Graph 3: The fundamental house price vs the actual house price in the United States(left) and the Netherlands (right)

The fundamental value of houses is calculated with the present value method, which is based on the work of black et al. (2006). Disparities between the actual and the fundamental values are from 2000 until 2006. Before the crisis in 2007, the actual values of house prices dropped

dramatically. Also, the fundamental value is decreased in 2007, but this is followed by a steady increase. The actual values of houses drop in the United States more than 20 percent in the period of the crisis.

In the Netherlands in 2000, the actual value of the house prices is higher than the fundamental value. Also, in the Netherlands there are disparities before the crisis in 2007. When the crisis hit the Netherlands, the housing prices decrease dramatically. Also, the fundamental value of houses stagnated for almost five years. After the new policy of quantitative easing of the European Central Bank in 2012, the market is recovering. The effects of this policy according to the work of Palley (2014), is that banks created more reserves, lowering interest rates and increase capital flows into the market. After this quantitative easing in 2012, the market is restoring, and from 2015 until 2018, the actual value increases rapidly.



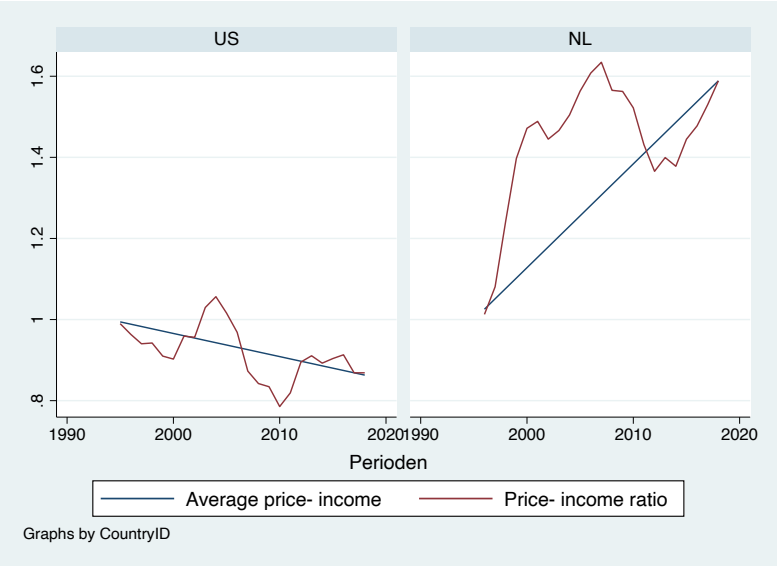
Graph 4: The difference between the actual and fundamental value for the United States and the Netherlands.

Graph 4 shows the differences between the actual and the fundamental values. According to the paper of Stiglitz (1990), are house prices which are departing from their fundamental value a risk for a bubble. According to the paper of Bourrassa et al. (2019), is there a criterion of twenty percent in order to calculate a bubble in the market. In the case of the United States, a housing bubble with the criterion of Bourassa et al. (2019) is it not the case, but before the crisis, the difference between the fundamental value and the actual value in the United States are more than fourteen percent. In the case of the Netherlands, the fundamental value and the actual value are not more than four percent departed from each other before the crisis. After the crisis the difference value between the fundamental and actual value are below zero, which indicates that the fundamental value of houses become bigger than the actual value. In 2016, the actual value become bigger than the fundamental value in the United States. Also, in the Netherlands the actual values are catching up with the fundamental value. The

steepness of the graphs is really high and, in this tense, could in the long run causing trouble if the twenty percent is reached or passed.

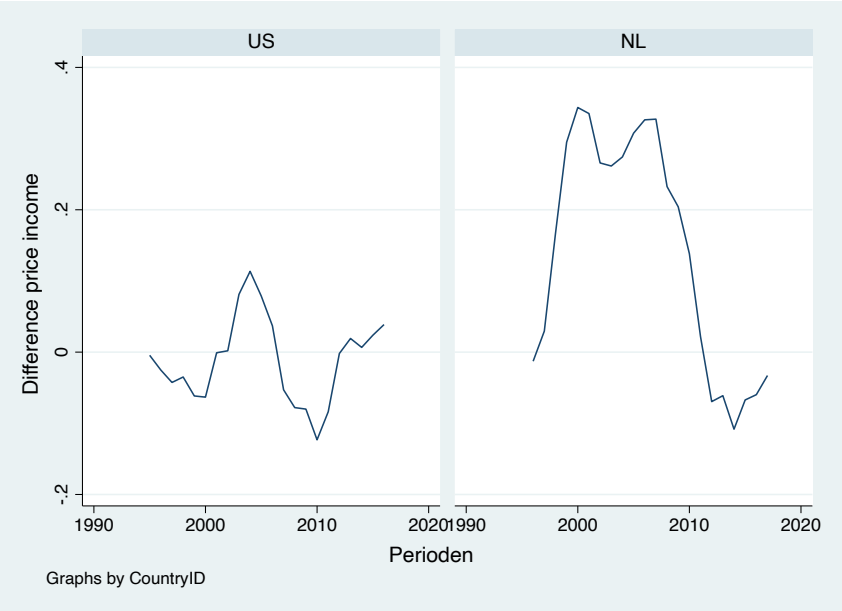
Results ratio measure analysis

After the results with regards to the regressions and the present value. This paragraph is testing whether the average price-income ratio in the Netherlands and the United States differ more than twenty percent of the long-term average. The price- income ratio is based on the yearly numbers of income and the average house prices in the region. According to the paper of Case and Shiller (2003), the criteria of twenty percent enough in order to calculate if a bubble is in the market.



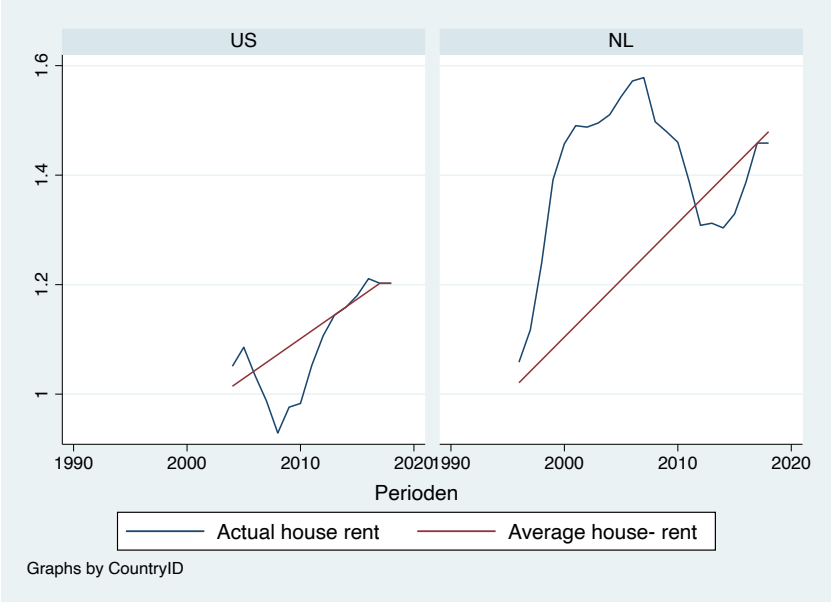
Graph 5 The price income ratio compared with the long-term average price income in the US and the Netherlands

First of all, the United States, has a decreasing average of the price income ratio. The reason is that income prices increase more than the house prices in the United States increase over time more than income. The average income increase was 4,5 percent between the period of 1995 and 2020. In the case of the Netherlands, house prices increased more over time than income. The average house price increased by 5 percent instead of income increase of almost 2,5 percent. This is followed by positive price-income ratio instead of the negative price-income ratio in the United States. If there is a housing bubble according to the paper of Case and Shiller (2003), the actual price- income ratio has to exceed the average price-income ratio by twenty percent. This is not the case for the United States, but this seems to be the case for the Netherlands. The next graph represents the results of the difference between the actual and the long-term difference.



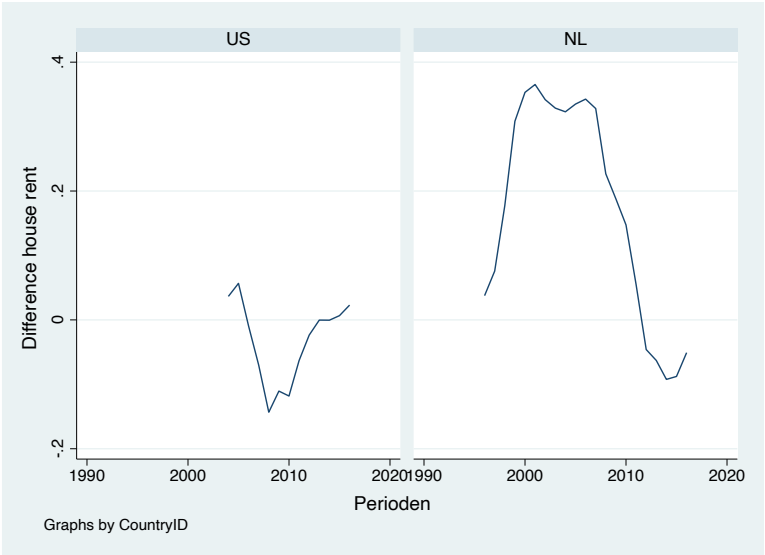
Graph 6 The differences between the price income ratio and long-term average price income in the US and the Netherlands.

Graph 6 shows the difference between actual and average price income value. The criteria of Case and Shiller (2003) is that the actual value exceeds the long-term average by twenty percent. As stated before, this is not the case in the United States. In the end of 90s, the house prices in the Netherlands increased by more than ten percent per year. Also, according to the IMF (2008) and graph 3, the house prices are overvalued, because the actual house prices rates growth more than the fundamental value. So, the price income ratio increased, which is shown in graph 6. Between the first quarter of 1998 and 2009, the price-income difference between the actual and the fundamental value is more than 20 percent. According to Case and Shiller (2003), a bubble could be predicted in this period. The difference between the United States and the Netherlands, is that the income and house prices in the U.S. almost in the same proportion growth.



Graph 7 *The actual house rent compared with the long-term average house rent in the US and the Netherlands*

Graph 7 indicates the ratio between the house prices and the yearly rent. For both countries there is a growth between this ratio. Especially, the Netherlands have positive growth between the house prices and the rent price. The average growth per year is almost two percent for the Netherlands, which is almost a half percent higher than the United States. Testing the hypothesis of both countries, a different graph is necessary. The next graph calculates the difference between the actual and the average value of the house price rent ratio.



Graph 8 *The differences between the actual house rent and long-term average house rent in the US and the Netherlands.*

According to the paper of Case and Shiller (2003) a twenty percent is the criterion in order to have a bubble in a region. In this case, the Netherlands have a bubble between 1998 and 2008. The United States doesn't have a housing bubble according to the criteria of Case and Shiller (2003). The reason for the more than twenty percent difference in the Netherlands is increased house prices in the late 90s and beginning of 2000. Rents prices in the Netherlands are by law controlled. So, this rate can't compete with the housing price increases. In the United States the rents and house prices are growing together, and therefore there is no gap between the actual and the average value.

Discussion

The last years, the housing market has made enormous booms and busts. According to the Parool (2020), the average house price in Amsterdam was for the first time above five hundred thousand euro. This was followed up by question if the market is heading into a housing price bubble. This thesis aims to provide answers which variables could possibly predict the house prices in a country. Some results are highly significant and influencing the house price. Also, literature is used in order to describe recent literature about the house prices, because according to black et al. (2006) the housing market doesn't receive many attentions before the crisis. According to black et al. (2006), the housing market is interesting for four reasons. The first reason is that the housing leads to more wealth effects than financial assets (Case et al.,2005; Benjamin et al., 2004). Second, according to the papers of Englund et al. (2002) and Flavin and Yamashita (2002) is housing one of the major assets of households. The third reason is that the effect of a housing bubble is twice as big as stock market bubble according to Helbling and Terrones (2003). The last reason is that buying a house not always is from an investment perspective.

Bourassa (2019) divided the types of methods part in to three parts, which is also done is this thesis. First, the present value method and the rental growth of black et al. (2006). The present value method tries to estimate the discount factor Campbell and Shiller (1988a, 1988b). The discount factor is used in order to discount future cash flows. The second method is the ratio measure is used. The ratios measures are the price rent ratio, the real median house price and the annual rent in an area. These formulas are based on the article of Bourassa et al. (2019). Also, the price- income ratio is used in this method. The last method is the regression method. The regression model is based on a logarithm function. Also, this model based their demand and supply on different variables.

This thesis is providing some literature in order to create a model, which could possible explain the house prices in a country. The model of Bourassa et al. (2019) is calculating the logged house price.

This model has as a dependent variable the logged house price index. In order to calculate the house price index, different variables are needed. These variables are unemployment, mortgage rate, consumer surplus, construction costs, term spread and the number of citizens.

First of all, the unemployment rate, which was according to the paper Abelson et al. (2005) expected to be negative. The results in this thesis reported that the unemployment rate was indeed negative and significant at a level of five percent. The unemployment rate has a negative relationship with the logged house price index. The second variable is the mortgage rate, which was also expected to be negative according to the paper of Abelson et al. (2005) and Bourassa et al. (2019). The mortgage rate coefficient was found negative and not statistically significant. So, the effect of mortgage rate on the housing prices could not be detected in this thesis. The third variable was consumer sentiment and was expected to be positive according to the paper of Johnson (2010). Also, consumer sentiment is affected through a vicious cycle, which is influencing the economic activities, but also influenced by the economic activities. Wilcox (2017), conclude that consumers based their short- and long-term consumption on future expectations. The results showed a different result than expected positive relationship. There was a small five percentage significant relationship between the consumer sentiment and the house price index. Also, with lags the relationship continues to be negative. So, is consumer sentiment is raising, the house price is decreasing. The fourth variable is the construction costs of buildings. This variable was expected to be positive according to the paper of Case and Shiller (1990), because construction costs influence the costs of new build homes. According to Capozza et al. (2002), the construction costs are highly influenced by new regulations and restriction costs. The expected positive relationship between constructions costs and the house price index was indeed a positive relationship with a one percent significant level. Thus, if the construction prices rise, the house price index becomes also higher, because new build homes are become more expensive. The fifth variable is the term spread and is calculated according to Bourrassa et al. (2019). The term spread was according Dotsey (1998), expected as positive related to the house price index, but the results reported no significant relationship. The sixth variable is income, which had some multicollinearity problems in the beginning. These problems are solved through the mean centering method according to the papers of Aiken and West (1991), Cohen and Cohen (1983), Jaccard, Wan and Turrisi (1990) and Smith and Sasaki (1979). The expected relationship was according to the papers of Abelson et al. (2005) and Tsatsaronis and Zhu (2004) positive. The results reported also a five percent significant positive relationship between income and the house price index. So, if the income rises in a country, the houses become more expensive. The last variable is the number of citizens, which had also some multicollinearity problems in the beginning which was also solved as described above. The expected relationship between the number of citizens and the house price index was expected to be positive according to Englund and Loannides (1997). The results reported a small negative effect; however,

this was not statistically significant, and might be explained by the relatively small amount of time that was researched in this thesis.

The second method was in order to test the present value. This method calculates the fundamental value of house prices. The fundamental value of houses is calculated for the Netherlands and for the United States. The actual value of houses is based on yearly data. According to the paper of Bourassa et al. (2019), the bubble criterion is when the actual value of houses are twenty ore more percent above the fundamental value. In both cases there was no twenty percent difference between the actual and the fundamental value. Still, in the United States was there a difference of fourteen percent before the crisis. Also, the difference between the actual value and the fundamental value become even negative in the crisis of 2007. After 2012 a positive trend occurs which creates a gap between the fundamental and the actual values. In this tense, the difference between the fundamental and the actual value become bigger and might even cause problems in the future.

The third method compares the price income ratio and the house rent ratio. The first method tested is the price income ratio for the Netherlands and the United States. The average price income ratio is calculated, and this number is compared with the actual value. According to the paper of Case and Shiller (2003), the criteria of twenty percent enough in order to calculate if a bubble is in the market. The United States has small positive difference before the crisis of 2007, but not enough to conclude that there is a bubble in the market. In the case of the Netherlands, there is a bubble in the market according to the criteria of Case and Shiller (2003), because the difference is more than twenty percent between 1998 and 2009. After the crisis, the bubble left, and the difference become even negative. During the crisis, the price income ratio become negative for both countries.

The second ratio is the house-rent ratio, which compared the house price with the costs of rents. Both countries have a positive growth for this ratio. In order to answer the hypothesis, the actual house rent ratio is compared with the average ratio. According to Case and Shiller (2003), a bubble could be identified when there more twenty percent between difference between the fundamental and actual value. According the criteria of Case and Shiller (2003) there is no housing price bubble in the United States in contrast to the Netherlands. There is a bubble between 1998 and 2008 according to the criteria of Case and Shiller (2003). The reason for this bubble seems to be an enormous housing price increases in the Netherlands in the late 90s and the first years of 2000.

There are some critical points in this thesis. The first point of critique is about the data set. The house price index of the United States is not based on the whole country. This data is derived from the FRED and is only based on different cities. The countryside of the United States is not included in the dataset. So, the house price index doesn’t reflect the entire nation.

The second point of critique is about the theories. In the literature, there has not been described a lot about the housing prices. However, the housing topic received more attention in the last years. So, there is less theory available than the classical subject in economics. According to the paper of Wheelock and Wohar (2009), authors of articles with regards to housing are more interested in analysis than in theory. So, authors are describing results and report effects, but they don't describe the theory behind. Also, black et al (2006) describes that the housing market doesn't receive a lot of attention before the crisis.

The last point of critique is about the high R^2, which is too high for a regression analysis. A high R^2 of 0.893 implies that 89.3 percent of the depending variable is accounted for. However, Bourrassa et al. (2019)| also described a high R^2. They also could not describe the reasons for this high number. This thesis is using the same variables as Bourrassa et al. (2019), because the same model was used.

In conclusion, the model of Bourrasa et al. (2019) is a helpful model in order to conduct a data analysis. The results showed variables which are significantly affecting the house price index. The important variables in order to predict the house price index are unemployment, construction costs and consumer sentiment. However, the data for the United States was not complete and doesn't reflect the whole country. Also, Black et al. (2006) conclude that the housing market was an upcoming topic in economics. Also, Wheelock and Wohar (2009) and Hornstein (2009) find that there is a lack of theory in economics for housing. This thesis has used some different methods as theory. The Bourrasa et al. (2019) model has been used for data analysis and found some significant variables but could not made prediction for a house price bubble. This paper indicates that there are some variables that affect the house price in the United States and the Netherlands. But in order to predict house price bubbles, there should be done more research and completer data collecting.

Appendix

1. Data acquisition of the various variables

	Netherlands	Amsterdam	US	New York
Prices of houses index	CBS (1995-2019) Monthly	CBS (1995-2019) yearly	FHFA 1980-2020 Yearly	FHFA 1980-2020 Yearly
Rent indexes	CBS (1959-2019) Yearly	CBS (2015-2019) yearly	FRED (1990-2019) monthly	Same
Consumer price indexes	CBS (1995-2019) Monthly	Same	Burea of Labor (BLS) 1887-2019	Same
Aggregate and per capita income	CBS (2011-2018) Yearly	CBS (2011-2018) Yearly	BEA (1959-2010) yearly BEA (2010-2019) Monthly	BEA (1959-2010) yearly BEA (2010-2019) Monthly
Income household	CBS (2011-2019) Yearly	CBS (2011-2019) Yearly	Factset (1995-2020) Monthly	Factset (1995-2020) Monthly
Gross domestic product	Factset /CBS (1995-2019) Quartely + yearly	Same Or only 2013	BEA (1969-2019)	2015-2019 yearly BEA (region
Population	CBS (1950-2019) yearly	CBS (1990-2019) yearly	BEA (1969-2019)	BEA (1969-2019)
Unemployment rates	CBS (2009-2019) Yearly	CBS (2010-2019) Yearly	BLS (1979-2019) yearly	BLS (1976-2019) monthly
Mortage interest rates	Dutch Central bank (2003-2019) Monthly	Same	Freddie Mac (1970-2019)	Same
Consumer sentiment indexes	CBS (1986-2019) Monthly	CBS (2017-2019) quarterly (region)	Michigan survey (1978-2019)	Same

Marginal income tax rates	Belastingdienst	Same	National bureau ec. research (1960-2019) yearly	National bureau ec. research (1977-2019) yearly
Capital income tax rate	Belastingdienst	Same		
Property tax rates	Included in maintenance	Included in maintenance	FRED	FRED
Maintenance costs	CBS (2012,2015,2018)	CBS (2012,2015,2018)	Harding, Rosenthal and Sirmans (2007)	Same
Construction costs indexed	CBS (2015-2019) Monthly	CBS (2015-2019) Monthly (region)	RS Means (1970-2020) yearly	Same

2. Rule of Thumb

Size of Correlation	Interpretation
0.90 to 1.00 (-0.90 to -1.00)	Very high positive (negative) correlation
0.70 to .90 (-0.70 to -0.90)	High positive (negative) correlation
0.50 to .70 (-0.50 to -0.70)	Moderate positive (negative) correlation
0.30 to .50 (-0.30 to -0.50)	Low positive (negative) correlation
0.00 to .30 (-0.00 to -0.30)	Negligible correlation

Hausman test

In this case, the zero hypothesis states that the random model is preferred. The alternative hypothesis is that the fixed model is preferred. The requirement for the p-value to be significant, is a value is a smaller than 0.05 in order to reject the zero hypothesis. The p-value is zero, which is smaller than 0.05. This means that the zero hypotheses are rejected, and the alternative hypothesis is accepted. In this analysis the fixed model will be used.

	Coef.
Chi-square test value	34.04
P-value	0.00

Table 2: Hausman (1978) specification test. This table shows the results of the Hausman (1978) specification test. The P value is 0.00, which means the alternative hypothesis is accepted and the fixed model should be used.

Breusch and Pagan test

If the variance of the Breusch and Pagan test is not equal to 0, the OLS regression approach is preferred above the panel data. Homoscedastic is a requirement for a regression analysis. The alternative hypothesis stated that the variance is not equal to zero, which indicated that there is heteroscedastic (Breusch and Pagan, 1979). Table 3 shows an overview of the Breusch and Pagan test.

	Var	sd = sqrt(Var)
LogHPIL	1361.763	36.90208
e	120.6954	10.98615
u	0	0

Test: Var(u) = 0
chibar2(01) = 0.00
Prob > chibar2 = 1.0000

Table 3: Breuschand Pagan test. The variance is 0 with a probability of 1.00.

Multicollinearity

```
. pwcorr logHPIL INCOME UNEM N R CS CONSTRUCT TS, star(0.05)
```

	logHPIL	INCOME	UNEM	N	R	CS	CONSTR~T	
logHPIL	1.0000							
INCOME	0.6497*	1.0000						
UNEM	-0.1183*	-0.1253	1.0000					
N	-0.0148	-0.8298*	0.1722*	1.0000				
R	-0.7740*	-0.7376*	-0.1278*	0.3570*	1.0000			
CS	-0.2324*	-0.0556	-0.7036*	-0.0692	0.2247*	1.0000		
CONSTRUCT	0.7528*	0.0471	-0.1423*	0.5538*	-0.4175*	-0.0957*	1.0000	
TS	-0.1906*	-0.2651	0.6217*	0.2309*	-0.0004	-0.4437*	-0.0814*	

	TS
TS	1.0000

Table 4. Correlation matrix which gives a visual representation of multiccollinearity. A value between one and minus one will be calculated. Value of 1 represents perfect correlation between two variables. The Asterisk indicates probability of <0.05. Respectively HPI (house price index), Income (income), UNEM (unemployment rate), N (number of citizens), R (mortgage rate), CS (consumer sentiment), TS (term spread).

Variable	VIF	1/VIF
INCOME	39.13	0.025554
N	29.06	0.034417
R	8.58	0.116609
CONSTRUCT	4.80	0.208533
UNEM	3.21	0.311082
CS	2.78	0.359873
TS	1.73	0.578156
Mean VIF	12.75	

Table 5: Variance inflation factor. This is a calculation of multicollinearity.

References

- Abelson, P., Joyeux, R., Milunovich, G., & Chung, D. (2005). Explaining house prices in Australia: 1970–2003. *Economic Record*, *81*, S96-S103.
- Ahrens, Ralf. "Predicting Recessions with Interest Rate Spreads: A Multicountry Regime-Switching Analysis." Journal of International Money and Finance, August 2002, 21(4), pp. 519-37.
- Abraham, J.M. and P.H. Hendershott (1993), "Patterns and Determinants of Metropolitan House Prices, 1977-1991," in L.E. Browne and E.S. Rosengren (eds.), Real Estate and the Credit Crunch, Federal Reserve Bank of Boston.
- Abraham, J.M. and P.H. Hendershott (1995), "Bubbles in Metropolitan Housing Markets" Journal of Housing Research. 6.
- Abraham, J. and P. Hendershott. 1996. Bubbles in Metropolitan Housing Markets. Journal of Housing Research 7(2): 191–207
- Aiken, Leona S., and Stephen G. West (1991), *Multiple Regression: Testing and Interpreting Interactions.* Thousand Oaks, CA, US: Sage Publications.
- Arnold, T., & Crack, T. F. (2004). Using the WACC to value real options. Financial Analysts Journal, 60(6), 78-82.
- El Alaoui, M., Bouri, E., & Azoury, N. (2020). The Determinants of the US Consumer Sentiment: Linear and Nonlinear Models. *International Journal of Financial Studies*, *8*(3), 38.
- Barnichon, Regis, Michael Elsby, Bart Hobijn, and Ayşegül Şahin. 2010. "Which Industries Are Shifting the Beveridge Curve?" FRBSF Working Paper 2010-32.
- Baker, D. (2008). The housing bubble and the financial crisis. *Real-world economics review*, *46*(20), 73-81.
- Bakshi, Gurdip S. and Zhiwu Chen, 1994, "Baby Boom, Population Aging, and Capital Markets," The Journal of Business, Vol. 67(2), pages 165-202.
- Benati, Luca and Goodhart, Charles. "Investigating Time-Variation in the Marginal Predictive Power of the Yield Spread." Journal of Economic Dynamics and Control, April 2008, 32(4), pp. 1236-72.
- Benjamin, J.D., P. Chinloy and G.D. Jud (2004), 'Real Estate versus Financial Wealth in Consumption', *Journal of Real Estate Finance and Economics*, Vol. 29, pp. 341–54.
- Bernard, Henri and Gerlach, Stefan. "Does the Term Structure Predict Recessions? The International Evidence." International Journal of Finance and Economics, July 1998, 3(3), pp. 195-215.
- Boelhouwer, P. (2017). The role of government and financial institutions during a housing market crisis: a case study of the Netherlands. *International Journal of Housing Policy*, *17*(4), 591-602.

- Bourassa, S.C., Hoesli, M. and Oikarinen, E. (2019), Measuring House Price Bubbles. Real Estate Economics, 47: 534-563. doi:10.1111/1540-6229.12154
- Black, A., P. Fraser and M. Hoesli. 2006. House Prices, Fundamentals and Bubbles. Journal of Business Finance and Accounting 33(9): 1535–1155.
- Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society*, 1287-1294.
- Canner, G., Dynan, K., & Passmore, W. (2002). Mortgage refinancing in 2001 and early 2002. *Fed. Res. Bull.*, *88*, 469.
- Campbell, J. and R. Shiller. 1988a. The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. Review of Financial Studies 1(3): 195–227
- Capozza, D. R., Hendershott, P. H., Mack, C., & Mayer, C. J. (2002). *Determinants of real house price dynamics* (No. w9262). National Bureau of Economic Research.
- Case, K.E. and R.J. Shiller (1988), "The Efficiency of the Market for Single Family Homes," American Economic Review, 79, 125-137.
- Case, K.E. and R.J. Shiller (1990), "Forecasting Prices and Excess Returns in the Housing Market," AREUEA Journal, 18, 253-273
- Case, K.E. J.M. Quigley and R.J. Shiller (2005), 'Comparing Wealth Effects: The Stock Market versus the Housing Market', *Advances in Macroeconomics*, Vol. 5, pp. 1–32.
- Case, K. & Fair, R. (2007). *Principles of Economics*. Upper Saddle River, NJ: Pearson Education. p. 54
- Chan, H.L., S.K. Lee and K.Y. Woo (2001), 'Detecting Rational Bubbles in the Residential Housing Markets of Hong Kong', *Economic Modelling*, Vol. 18, pp. 61–73.
- Cohen, Jacob and Patricia Cohen (1983), *AppliedMultiple Regression/Correlation Analysis for the Behavioral Sciences.* (2nd ed.) , NJ, Hillsdale: Lawrence Earlbaum Associates
- Cubillas, E., & González, F. (2014). Financial liberalization and bank risk-taking: International evidence. *Journal of Financial Stability*, *11*, 32-48.
- Cuthbertson, K. (1996), *Quantitative Financial Economics* (John Wiley & Sons, New York, USA)
- De Kruijf, 2020. Over 17.4 Over 17.4 million inhabitants in the Netherlands. CBS. *Derived from https://www.cbs.nl/en-gb/news/2020/01/over-17-4-million-inhabitants-in-the-netherlands*
- Dissanaike, G. (1997). Do stock market investors overreact? *Journal of Business Finance & Accounting*, *24*(1), 27-50.
- Dotsey, M. (1998). The predictive content of the interest rate term spread for future economic growth. *FRB Richmond Economic Quarterly*, *84*(3), 31-51.
- Dua, Pami, 2008, "Analysis of Consumers' Perceptions of Buying Conditions for Houses," Journal of Real Estate Finance and Economics, Vol. 37, pages 335-350

- Dynan, K., Elmendorf, D., & Sichel, D. (2012). The evolution of household income volatility. *The BE Journal of Economic Analysis & Policy*, *12*(2).
- Englund, P., & Ioannides, Y. M. (1997). House price dynamics: an international empirical perspective. *Journal of housing economics*, *6*(2), 119-136.
- Englund, P., M. Hwang and J.M. Quigley (2002), 'Hedging Housing Risk', Journal of Real Estate *Finance and Economics*, Vol. 24, pp. 167–200.
- Estrella, Arturo and Hardouvelis, Gikas A. "The Term Structure as a Predictor of Real Economic Activity." Journal of Finance, June 1991, 46(2), pp. 555-76.
- van Giesen, Roxanne I., and Rik Pieters. 2019. Climbing out of an economic crisis: A cycle of consumer sentiment and personal stress. Journal of Economic Psychology 70: 109–24
- Estrella, Arturo, and Gikas A. Hardouvelis. "The Term Structure as a Predictor of Real Economic Activity," Journal of Finance, vol. 46 (June 1991), pp. 555–76.
- Estrella, Arturo, and Frederic S. Mishkin. "The Predictive Power of the Term Structure of Interest Rates in Europe and the United States: Implications for the European Central Bank," European Economic Review, vol. 41 (July 1997), pp. 1375–1401.
- Flavin, M., & Yamashita, T. (2002). Owner-occupied housing and the composition of the household portfolio. *American Economic Review*, *92*(1), 345-362.
- Fu, X., Lin, Y., Molyneux, P., 2014. Bank competition and financial stability in Asia Pacific. J. Bank. Financ. 38 (January), 64–77
- Froot, K. A., & Obstfeld, M. (1991). Exchange-rate dynamics under stochastic regime shifts: A unified approach. *Journal of International Economics*, *31*(3-4), 203-229.
- Haffner, M. E., & Boumeester, H. J. (2010). The affordability of housing in the Netherlands: An increasing income gap between renting and owning?. *Housing Studies*, *25*(6), 799-820.
- Hair, J. F. Jr., Anderson, R. E., Tatham, R. L. & Black, W. C. (1995). *Multivariate Data Analysis*, 3rd edn. New York: Macmillan.
- Harris, R., & Meir, A. (2015). Recourse structure of mortgages: a comparison between the US and Europe. *CESifo DICE Report*, *13*(4), 15-22.
- Helbling, T., & Terrones, M. (2003). When bubbles burst. *World Economic Outlook*, *4*(2), 61-94.
- Helbling, T., & Terrones, M. (2003). Real and financial effects of bursting asset price bubbles. *IMF World Economic Outlook*, *61*, 94.
- Herring, R. J. (2007). The rocky road to implementation of Basel II in the United States. *Atlantic Economic Journal*, *35*(4), 411-429.
- Himmelberg, C., C. Mayer and T. Sinai. 2005. Assessing High House Prices: Bubbles, Fundamentals and Misperceptions. Journal of Economic Perspectives 19(4): 67–92.
- Hinkle D.E, Wiersma W, Jurs SG. Applied Statistics for the Behavioral Sciences. 5th ed. Boston: Houghton Mifflin; 2003.

- Houle, J., & Berger, L. (2014). Is student loan debt discouraging home buying among young adults. *Association for Public Policy and Management.*
- Hornstein, A. (2009). Problems for a fundamental theory of house prices. *FRB Richmond Economic Quarterly*, *95*(1), 1-24.
- Jaccard, James, Choi K. Wan, and Robert Turrisi (1990), ‘ The Detection and Interpretation of Interaction Effects BetweenContinious Variables in Multiple Regression,’ *Multivariate Behavioral Research, 25, 4, 467-78*
- Jacobsen, D. H., & Naug, B. E. (2005). What drives house prices?
- Johnson, M. A. (2010). Studying How Changes in Consumer Sentiment Impact the Stock Markets and the Housing Markets.
- Lahiri, Kajal, and Yongchen Zhao. 2016. Determinants of consumer sentiment over business cycles: Evidence from the US surveys of consumers. Journal of Business Cycle Research 12: 187–215.
- Lee, C.-C., Hsieh, M.-F., 2014. Bank reforms, foreign ownership, and financial stability. J. Int. Money Financ. 40 (February), 204–224
- Lind, H. 2009. Price Bubbles in Housing Markets: Concept, Theory and Indicators. International Journal of Housing Markets and Analysis 2(1): 78–90.
- Mayer, Christopher and Tsur Somerville 2000. “Land Use Regulation and New Construction?" Regional Science and Urban Economics, 30: 639-662.
- Meen, G. (1996). Spatial aggregation, spatial dependence and predictability in the UK housing market. *Housing studies*, *11*(3), 345-372.
- Meen, G. (1996), ‘Ten Propositions in UK Housing Macroeconomics: An Overview of the 1980s and Early 1990s’, *Urban Studies*, Vol. 33, pp. 425–44.
- Merton, R. 1973. An Intertemporal Capital Asset Pricing Model. Econometrica 41(5): 867–888.
- Merton, R 1980. On Estimating the Expected Return on the Market: An Exploratory Investigation. Journal of Financial Economics 8(4): 323–361
- Mian, Atif R. and Sufi, Amir, 2009. “The Consequences of Mortgage Credit Expansion:Evidence from the U.S. Mortgage Default Crisis”, Quarterly Journal of Economics.
- Mian, A., & Sufi, A. (2011). House prices, home equity-based borrowing, and the US household leverage crisis. American Economic Review, 101(5), 2132-56.
- Mutl, J., & Pfaffermayr, M. (2011). The Hausman test in a Cliff and Ord panel model. *The Econometrics Journal*, *14*(1), 48-76.
- Neter, J., Wasserman, W. & Kutner, M. H. (1989). Applied Linear Regression Models. Homewood, IL Richard D. Irwin.
- O’brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*(5), 673-690.

- Palley, T. I. (2014). Monetary policy in the US and EU after quantitative easing: the case for asset-based reserve requirements (ABRR). *Real World Economic Review*, *68*, 2-9.
- Roche, M. J. (2001). The rise in house prices in Dublin: bubble, fad or just fundamentals. *Economic Modelling*, *18*(2), 281-295.
- Roehner, B. M. (1999). Spatial analysis of real estate price bubbles: Paris, 1984–1993. *Regional science and urban economics*, *29*(1), 73-88.
- Rotemberg, Julio J., 2008, "Subprime Meltdown: American Housing and Global Financial Turmoil," Harvard Business School Case Number 708-042, pages 1-21.
- Sanders, A. (2008). The subprime crisis and its role in the financial crisis. *Journal of Housing Economics*, *17*(4), 254-261.
- Smith, Kent W. and M. S. Sasaki (1979), 'Decreasing Multicollineairity: a Method for Models with Multiplicative Functions,' *Sociological Methods and Research,* 8 (1), 35-56
- Stiglitz, J.E. 1990. Symposium on Bubbles. Journal of Economic Perspectives 4(2): 13–18.
- Stock, James H., and Mark W. Watson. "New Indexes of Coincident and Leading Economic Indicators," in Olivier J. Blanchard and Stanley Fischer, eds., NBER Macroeconomics Annual 1989. Cambridge, Mass.: MIT Press, 1989, pp. 352–94
- Tsatsaronis, K. and Zhu, H. (2004), 'What Drives Housing Price Dynamics: Cross-Country Evidence', BIS Quarterly Review, March, 65–78.
- Vélez-Pareja, I., Ibragimov, R., & Tham, J. (2008). Constant leverage and constant cost of capital: A common knowledge half-truth. *Estudios Gerenciales*, *24*(107), 13-33.
- Weidner, J., & Williams, J. C. (2011). What is the new normal unemployment rate?. *FRBSF Economic Letter*, *5*.
- Wheelock, D. C., & Wohar, M. E. (2009). Can the term spread predict output growth and recessions? A survey of the literature. *Federal Reserve Bank of St. Louis Review*, *91*(5 Part 1), 419-440.
- Wilcox, J. A. (2017). Constructing the Home Purchase Sentiment Index. *FRBSF Economic Letter*, *2017*, 14.
- Zhou, W. and D. Sornette. 2006. Is There a Real Estate Bubble in the US? Physica A 361(1): 297–30