A New Approach in the Behavior of House Prices
-A Study on the Conditional Heteroskedastic Effects in House Price Functions

Master thesis in economics, 15 ECTS
Department of Economics
Lund University, School of Economics and Management

Author: Henrik Björk
Date: May 27, 2013

Supervisor: Fredrik N G Andersson
Abstract

House prices and their movements are a topic of much interest in today’s society. Most of the population in a country is affected by house price movements in one way or another. The debate mostly focuses on house price bubbles. In this thesis it is argued that the debate should widen its focus. During times of crises, house prices’ conditional variance increases and this harms the economy. Thus it is of importance to model this behavior. Research about house prices conditional variance has so far only been concentrated on the volatility of that single variable. In this thesis, it is argued that other fundamental variables that explain house prices should be added to this analysis. It is the conditional variance of the residuals of a house price function that should be analyzed rather than only the house price itself. Conditional heteroskedastic effects are found and modeled for the residuals from a house price function. It is concluded that both the explanatory power and the fit of the model increase by adding those variables in the analysis.

Keywords: House Prices, Conditional variance, ARCH models
# Contents

1. Introduction ...........................................................................................................................................3
2. Theory ......................................................................................................................................................6
3. Data ...........................................................................................................................................................9
4. Method .....................................................................................................................................................15
   4.1 ARCH and GARCH models ...................................................................................................................16
   4.2 TARCH and EGARCH models ..............................................................................................................18
   4.3 Testing for ARCH and GARCH effects and how to specify a model .............................................20
5. Results ....................................................................................................................................................21
   5.1 Cointegration results ............................................................................................................................21
   5.2 Regression results .................................................................................................................................25
6. Concluding remarks ...............................................................................................................................34
7. References .................................................................................................................................................36
8. Appendix ..................................................................................................................................................39
1. Introduction

In the latest report by the Swedish Fiscal Policy Council, Jonung et al. (2013) conclude that Sweden might be experiencing a house price bubble. The situation is not yet severe, with house prices probably overvalued by 15-20%. Still they urge the Swedish government to take actions against the problematic situation. House prices and private borrowing have been increasing for some time and the bursting of a bubble could be contagious to other parts of the economy. Hott & Jokipii, (2012) conclude that the link between low interest rates and house price bubbles is strong. They argue that it is dangerous to have an interest rate set too low for a longer period of time. In Sweden, the interest rates have been set low and remained low for a quite long time. This fact is another warning for the Swedish housing market and thus for the Swedish economy. If Sweden, or any other country, is experiencing a house price bubble it is of great importance to understand the behavior of house prices if the bubble bursts. Not least because of the possibility for housing market disturbances to become contagious and impact the rest of the economy.

In the long run, house prices are usually well described by fluctuations around a function of fundamental variables in the economy. A house price bubble is commonly defined as a situation where such variables do not explain house prices very well. McMillan & Speight (2010) speak of house price bubbles as caused by the presence of a non-fundamental component. For some other reason the house price is not what it should be. It could be driven by expectations. Fluctuations around the function are normal. However, when they become unusually large they are likely to cause problems in the economy. Then the question arises, how do house prices behave when they start to deviate more than usual from the function of fundamental variables? Do these deviations lead to further deviations? And how long does it take for the deviations to come back to more normal levels? Remarkably little research has been done in this area.

These questions need to be answered not only in Sweden but in many other countries. In the last decades house price movements have been more of current interest in debates. Increased interest is found both in media and in the academic world. The debate is often about the country’s risk of experiencing a house price bubble. In my point of view, it should widen its focus and also deal with the above mentioned questions. House price bubbles obviously
appear from time to time. If they cannot easily be avoided, there should also be a debate on how to dampen the effects when the bubble bursts. If policy makers want to reduce those effects, they must first know how house prices behave when a bubble bursts. Thus in order for them to make the right decisions, more research is needed in this area. The deviations from the house price function need to be modeled.

For Sweden, house prices have kept rising even though the country was hit by the latest financial crisis at the end of 2007 (Statistics Sweden). For USA and UK, an answer to the questions above, would have been preferable before the latest financial crisis. During the crisis both of these countries experienced an unusually rapid fall in house prices (Freddie Mac and Halifax). The questions are very hard to answer, even for the most skilled and educated economists. There are always economists with different answers and opinions. Miller & Peng (2006) concludes that volatile house prices have contagious effects on the rest of the economy. This conclusion illuminates the importance of not just to determine if a country is in a house price bubble, but also answering the above questions.

The central banks set their interest rate. It is probably the most effective tool to affect the movements on the housing market and thus also private indebtedness. Therefore housing price development and the methodology for deciding whether a country is in a house price bubble should be of great concern for central banks. It is also important to know what to expect in situations of volatile and unexplainable house price movements. The aim of this thesis is to find out if there are more concerns that the central banks, mortgage lenders and house buyers should take into account in their analysis of house markets.

From what has been mentioned so far, one can conclude that the variance of house prices is not constant over time. During times of crises, the volatility usually increases. This fact raises the issue of whether the variance in the deviation from the house price function also shows this pattern. I then decided to further investigate this issue by trying to answer the question: *Are there conditional heteroskedastic effects in the deviations from the fundamental function that explains house prices over the long run?* If that is the case, it implies that the fluctuations around the house price functions increase its amplitude during unstable times, which is likely to cause instability in the rest of the economy. Then central banks should take this into account in their house market analysis. For mortgage lenders, like commercial banks, these effects should also be taken into account. In economics volatility is usually a measurement of
risk. If the mortgage lenders do not take the conditional heteroskedastic effects in account, they fail to identify the true risk of a house loan. The private sectors biggest cost is usually housing. Thus the volatility of the housing market could be seen as a measurement of the vulnerability of the whole economy. This in combination with the existing levels of lending and interest rates too. If those effects do exist, an attempt will be made to estimate them. The best way to model conditional heteroskedastic effects is by an autoregressive conditional heteroskedastic model (ARCH) or a generalized ARCH model (GARCH).

The possibility of conditional heteroskedastic effects is investigated in three countries. The countries are Sweden in 1986-2012, the USA in 1975-2012 and the UK in 1983-2012. If those effects are detected, they will be captured by estimating ARCH or GARCH models. Then it should fit with the house price movements in a better way than when ignoring these effects. If there is conditional heteroskedasticity, those effects should be included in the methodology for answering the questions posed earlier. The focus of this thesis is to answer the question about conditional heteroskedasticity, not to determine whether a country is in a house price bubble or not.

So far, nobody has tried to answer this question. The house price functions and house price bubbles spoken of earlier, are well researched areas, see e.g. Clausen et al. (2011), and Hott & Jokipii (2012). Possible conditional heteroskedastic effects appear to not be taken into account in the literature. However, some studies with different type of ARCH and GARCH models for house prices do exist. The authors point out that even though this is such an important subject, very little research has been done in this area. In those studies, it is only the conditional heteroskedastic effects in the house price itself that is the focus. Thus they try to find those effects in the residuals from an ARMA model of house prices. In this thesis, those effects are examined in the residuals of a house price function of fundamental variables. This should be a better way of explaining the behavior of house prices.

Tsai et al. (2010) estimate different types of AR(1)-ARCH models on the house prices for old and new houses in the UK in 1953-2007. They estimate a mean function where the house price is an autoregressive function of order one, i.e. a function that depends on its first lag. Furthermore they model different types of ARCH models from the residuals. When they test for conditional heteroskedasticity, they do find such effects. Their different kinds of ARCH models are also significant on conventional levels. This indicates a variance that does vary
over time. They also find different states of volatility and therefore propose a regime-shifting ARCH model (SWARCH) for best describing the conditional variance of the house prices.

Furthermore, many researchers focus on the behavior of the housing markets and compare them to financial markets. Karaglou et al. (2011) use American data from the period 1987-2009 for a number of American cities. They estimate GARCH in mean models, which are models for identifying the risk premium of an asset. They find some evidence that house prices behave like other assets. Both the risk-return relationship and the asymmetric shock behavior (i.e. negative shocks have greater impact on the conditional variance) are the same for houses as for other assets. Not for all cities though. They argue that the financial sector has failed to account for the riskiness of house prices over the last 20 years. The authors identify structural breaks in their data due to different economic events and place strong emphasis on this issue. The latest financial crisis is an example of such an event. Morley & Thomas (2011) perform a similar study but on data from UK in the period 1972-2008. Their results and conclusions are in line with those described by Karaglou et al. (2011). They find some evidence that house prices behave like equities and commodity prices in the risk-return relationship and asymmetric effects. However, the evidence they find for asymmetric effects is that positive shocks have greater impact on the conditional variance.

Miles (2009) looks at data from a number of American metropolitan areas in the period 1975-2008. He argues that house prices often display very high persistence, which standard GARCH models are not able to capture. Instead, he suggests that one should use a component GARCH model (CGARCH) when estimating the conditional variance for house prices in the cases with very high persistence. In the latter three studies, they all use different types of ARMA models as mean functions in their estimation of different ARCH and GARCH models. Miller & Peng (2006) estimates GARCH and panel VAR models on American data in the period 1990-2002 to conclude that the volatility in house prices affects other variables growth rates and volatility in the economy. They conclude that conditional volatility of the house prices does affect other economic variables.

2. Theory

The residuals analyzed in the estimation of the ARCH and GARCH models, come from a house price function. As mentioned in the introduction, much research has been done about
these kinds of fundamental functions. From the background of this research and the accessible
data for the three countries I choose which variables to include. All variables are shown in
Equation (1).

$$\ln(p_t) = \beta_0 + \beta_1 \ln(GDP_t) + \beta_2 \ln(K_t) + \beta_3 \ln(pop_t) + \beta_4 r_t + \epsilon_t,$$

where $p_t$ is real house prices, $GDP_t$ is real GDP, $K_t$ is the real capital stock of housing, $pop_t$ is
the population, $r_t$ is the real interest rate and $\epsilon_t$ is the residual used in the estimation of ARCH
and GARCH models. If the fit of the model improves by additional lags on the right hand
side, those will be included. Equation (1) can be seen as a benchmark with the possibility to
add more lags. In this decision the change in Akaike and Schwarz Bayesian information
criteria (AIC and SBC) will used as guidelines.

In Equation (1), $GDP_t$ is a proxy for the wealth in a nation and should thus have an effect on
house prices. It is reasonable to assume that greater wealth leads to increased house prices.
Therefore this relationship is expected to be positive. The interest rate is the price of
borrowing money. Therefore it should have large emphasis in explaining house prices. If the
interest rate goes up, it is more expensive to borrow and thus house prices should decrease. It
is then reasonable to expect a negative relationship between the two. Clausen et al. (2011)
also use variables for wealth and the interest rate. They also argue that house prices are
determined by demand and supply. From that argument, the conclusion can be drawn that the
capital stock in housing and the population should matter in the determination of house prices.
An increased capital stock of housing implies an increased supply. If the supply goes up it
makes sense to expect a fall in house prices. The sign of the coefficient is expected to be
negative. On the other hand one can also imagine the possibility of a rise in the capital stock
of housing when house prices increase. The higher price gives more profit opportunities for
construction companies companies. Then the relationship is expected to be positive.

In order for $\epsilon_t$ to be stationary there are two possibilities. One is that all variables are
integrated of order zero I(0). The integration order is the number of times a variable has to be
differentiated in order to be stationary. The other possibility is that all variables are not I(0)
but they are integrated of the same order, I(d), and the variables are cointegrated. Usually in
macroeconomic time series data, the variables are integrated of order one, I(1). Then the
variables have to be cointegrated in order to avoid a spurious regression (Enders 2010)).
When $\varepsilon_t$ is ensured to be stationary it is time to analyze its conditional variance. If the conditional variance varies over time, it implies that the performance in Equation (1) is different from time to time. In times of house price bubbles, the performance of Equation (1) is weaker than usual. For other reasons the house prices then starts to deviate more than usual and thus further aspects are unexplained by the fundamental variables. This is most likely due to speculative behavior in housing markets, a severe situation that often leads to a worsening of other macroeconomic variables in a country. This could be increased interest rates, decreased GDP and increased unemployment.

Increased volatility often tends to create even higher volatility before it finally ebbs away. If the conditional variance is assumed to last for a long time, a GARCH model is better to use than an ARCH model. The latter only explains the $p$ chosen lags’ direct effects on the conditional variance. In addition, a GARCH model includes at least one lag of the conditional variance. Thus it captures the effects on the conditional variance from all previous lags. From an economic point of view, it is preferable to find out how long the increased conditional variance is expected to last. The longer it lasts, the greater impact it will have on the economy and the more important it is for policy makers to take action to dampen those effects. Thus a GARCH model or a high order ARCH model is preferred for explaining the conditional variance in that case. From an economic point of view, it is also interesting to find out if the increased conditional volatility in period $t$ is created by an increased volatility in period $t-1$. Then unusually big shocks can be self-fulfilling and create greater amplitude than necessary. This is a fact important for policy makers to be aware of.

There could be asymmetric effects in $\varepsilon_t$. Asymmetry implies that positive and negative shocks of the same size do not affect the conditional variance equally. The most common case is that negative shocks create higher conditional variance than a positive shock of equal size, so called leverage effects. Obviously negative shocks then have an effect of creating panic. This case is common in different kinds of asset markets, see e.g. Nuno et al. (2007) and Goudarzi & Ramanarayanan (2011). As mentioned in the introduction, these effects have also been detected for house prices. If that really is the case, then unusual falls in house prices tend to incite panic behavior.
3. Data

Sweden’s house prices contain two series, one on quarterly data between 1986Q1-2012Q4 and one on monthly data between 1999M3-2012M12. The source is in both cases Statistics Sweden. The series are nominal and thus they are deflated by CPI to present them in real terms. The reason of the inclusion of two series for Sweden is that the monthly data is very short. Monthly data is preferable because otherwise important information might be lost. Sweden’s latest dramatic fall in house prices was during its financial crisis of the early 90s. Therefore this period should be important to include, which the monthly series does not. Since the conditional variance of the deviations from the house price function is analyzed in this thesis it is especially important. This time period is likely to be a source for detecting such effects. For the USA the Freddie Mac House Price Index is used. This series is given on a monthly basis and stretches from 1975M1 to 2012M12. The data for UK is brought from Halifax and is also given on a monthly basis. It stretches from 1983M1 to 2012 M12. The latter two series are also nominal, so to provide them in real terms they are deflated by CPI. Furthermore seasonally adjustments are made and the natural logarithm of all house prices is taken before the regressions are run.

The data on GDP comes from OECD for all countries. All GDP series cover each county’s time interval for which house prices are found. They are given on a quarterly basis and in real values. The American data is in year 2005’s prices, the British data is in year 2009’s prices and the Swedish data is in year 2011’s prices. The data has to be on a monthly basis to fit with the house prices. All GDP series are therefore transformed by cubic-match last interpolating method in the econometric software E-views 7. Then seasonal adjustments are made and the natural logarithm is taken for all values. GDP is the variable for wealth in a country. There are better variables for measuring wealth, like disposable income and the private holdings of financial assets used by Clausen et al. (2011). However, that kind of data is not so easily found for all countries and time periods. GDP is probably correlated with those variables and hence it should be a good variable for capturing the wealth in a country.

For the capital stock of housing, the transformation process of the data is a bit more complicated. As a proxy for housing investments gross fixed capital formation in housing is used. The data is collected from OECD and is given on a quarterly basis in real values with the same base years as for GDP. The series for the USA and Sweden start as early as 1963Q1
and for the UK it starts in 1966Q1. All three of them stop at 2012Q4. From the GFCF in housing the capital stock in housing is calculated as early as the data allows. It is calculated by the perpetual inventory method. For more details about this method, see for instance Berlemann & Wesselhöft (2012). Once the series for the capital stock of housing is calculated they are transformed in the same way as GDP.

The population data comes from the World Bank’s database. These series are given on a yearly basis. The first transformation is by interpolation to provide them on a monthly basis and then the natural logarithms are taken. One might question the transformation of a series from yearly to monthly basis, but in this particular case I see no problems with that transformation. The population in developed countries usually grows at a steady and slow pace. It is reasonable to assume that population growth does not fluctuate much during a single year.

The interest rates used are the ones for government bonds with 10 years of maturity for each country. This data also comes from the OECD and are on a quarterly basis. Before the inclusion into the regression, the interest rates are interpolated just like in the previous cases. The series are in nominal values and hence they are deflated by CPI to provide them in real terms. The logarithm of the interest rates is not taken because the change in percentage units of interest rates is more intuitive than percentage change. It would be preferable to acquire data on the interest rates set by commercial banks on their house loans. Since that kind of data is not easily found, the government bond interest rates are used as proxies. Different kinds of interest rates tend to follow each other and thus it should be a good proxy.

All of the above described variables are illustrated in Figure 3.1 below. In the figure the variables are shown in real values, before any adjustments for seasonality are made and before taking the natural logarithm of the values. All series’ but the ones for the interest rates are also transformed into index form with the first observation set to 100. Then it is easier to see how cointegrated the variables in a single country are. It is also easier to compare the variables between different countries. There are five panels in Figure 3.1. The interest rates are shown in a separate panel, Panel (e). The other four panels contain all other variables. One for the USA, Panel (a), one for the UK, Panel (b), and two panels for Sweden, one on a quarterly basis, Panel (c) and one on a monthly basis, Panel (d).
Figure 3.1: The variables development over time

Panel (a): All variables but the interest rate for USA in 1975M01-2012M12

Panel (a): All variables but the interest rate for UK in 1983M01-2012M12

Source: Freddie Mac, OECD, World Bank

Source: Halifax, OECD, World Bank
Panel (c): All variables but the interest rate for Sweden in 1986Q1-2012Q4

Panel (d): All variables but the interest rate for Sweden in 1999M03-2012M12

Sweden (Q)

Source: Statistics Sweden, OECD, World Bank

Sweden (M)

Source: Statistics Sweden, OECD, World Bank
Panel (e): Interest rates for all countries in the time intervals as above for each country

Interest rates for all countries

Source: OECD
Even though the house price situation in the USA is very much debatable, one can conclude from Figure 3.1 that house prices in the USA have not experienced peaks of house price growth nearly as much as in the UK and Sweden. The highest peak during these time periods was in the second half of 2006. Then house prices were about 70% higher than the levels in the beginning of 1975. This can be compared with the peak in the UK in early 2007, where house prices had went up by about 220% since the level in 1983, and Sweden’s peak in late 2010 where house prices had risen by about 150% since the level in 1986. The house prices in USA move in cycles from 1975 to 2000 and then they continue to grow at rapid pace. Englund (2011) concludes that a cyclical pattern is a common behavior for house prices in most countries. The rapid increase stops just before the latest financial crisis in the world and then the heavy fall takes off. This was the trigger of the latest crisis according to many researchers, Murthy et al. (2009) being some of them.

The capital stock of housing has grown much in the USA. When it reached its peak in late 2007, it had risen by about 130% since 1975. The housing stock follows a similar path to house prices, but with a lag. Its development is smoother than house prices though, which is intuitive since investments are usually big decisions and stretch over long periods of time. GDP also seems to follow house prices quite well. An interesting fact about GDP and its movement with house prices is that every time house prices start to fall from a peak, it is followed by a fall in GDP. The fall in GDP always comes after the fall in house prices. The population grows at a quite constant pace. It does not really follow the house price movements, but the trend is upwards, which it also is for house prices.

In the UK, house prices fluctuate more over time and have experienced tremendous growth over the period. Even after the latest rapid fall, the house price level is still about 120% higher than in 1983. Also in the UK, house prices move in a cyclical pattern but with a more upwards trend than in the USA. The UK has experienced two peaks followed by two heavy falls during this time interval. The first was during the crisis of the early 1990s and the second was during the latest financial crisis. Also in this case, house prices start to fall before GDP does and the capital stock of housing tends to follow house prices with a lag even though it is in a much smoother way. When it comes to population growth, one can apply the same reasoning as for the USA.
Sweden is different from the other countries in many aspects. During the first half of the sample, house prices exhibit a cyclical pattern like for the other two countries. House prices grow until the crisis of the early 1990s and then experience a heavy fall. During the second half of the sample they behave different. The trend is now that house prices keep growing. There is a small fall in connection with the latest financial crisis, but it does not take a long time before house prices start growing again. Also in Sweden, house prices seem to fall before GDP does. However, this pattern is not as clear-cut as for the two previous cases. In connection with the crisis during the 1990s house prices do fall before GDP but GDP actually starts to fall before the fall becomes unusually fast. As far as the latest financial crisis is concerned house prices start to fall before GDP does, but house prices do not fall very much compared to the other countries. From the end of 2007 and onwards the movements of GDP and house prices are almost identical. The capital stock of housing for Sweden is completely different from the previous cases. In the first fourth of the sample it seems to follow house prices with a lag, but then it falls and stays at a lower level than in 1986 for the whole sample. When house prices keep growing for some time, the housing stock starts to grow a little again. The population growth looks similar for all three countries.

The interest rates for all three countries follow each other quite well. Throughout the sample Sweden’s interest rate is the highest, the USA’s interest is the lowest and the UK’s interest rate is somewhere between them. At the end of the 1990s the interest rates for all countries reaches a level of about four percent and the trend now indicates faster falling interest rates (with the exception of the USA for some time during the crisis). Maybe the low interest rates were a key factor for the rapidly increasing house prices in all three countries. House prices in all three countries start to grow fast in this specific time period. Interest rates in all three countries continue to go down after falls in house prices, but that could rather be a consequence of the fact that all countries have not come through the latest financial crisis. This induces central banks to keep their interest rates low.

4. Method

Usually in econometrics one would like the variance of the residuals to be constant and independent of time. This is actually one of the Gauss-Markov assumptions for OLS to be the best linear unbiased estimator according to Veerbeck (2012). If the variance of the residuals is not constant, one usually tries to correct for it. This is usually made by GLS or a robust
heteroskedastic estimator. The variance is often a measurement of risk in economics. If the variance is assumed to be constant when it is not, one actually fails to describe the true riskiness of an asset in a certain time period. Obvious cases of this description are the house markets described by the researchers mentioned in the introduction. They conclude that the conditional variance of house prices and thus also the riskiness of owning a house, is very different in different periods of time. Large shocks tend to create a period of greater volatility. This pattern is very common in financial economics and hence ARCH and GARCH models are mostly used in financial economics. In this thesis, those models are the most important tool in the method.

4.1 ARCH and GARCH models

In order to be able to model these kinds of effects, Engle (1982) developed the ARCH model. Later Bollerslev (1986) extended Engle’s paper and introduced the GARCH model. Today, so many extensions on these models have been made that ARCH and GARCH models have become a group of many models. Two models commonly used for estimating asymmetric effects are the threshold ARCH model (TARCH) originally developed by Glosten et al. (1994) and the exponential GARCH model (EGARCH) originally developed by Nelson (1991). All these models are also well described by Enders (2010).

An ARCH (p) model takes the form of Equation (2), where the conditional variance in time t, \( \sigma_t^2 \), is determined by an intercept, \( \alpha_0 \), and the sum of the squared residuals in the previous periods times their coefficients until period t-p. The restrictions are that the intercept, \( \alpha_0 \), and all \( \alpha_i \)'s must be positive in order to ensure a positive variance. The sum of the \( \alpha_i \)'s must be less than one in order for the model to be stationary.

\[
\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2
\]  

(2)

A GARCH (p, q) model takes the form of Equation (3). The conditional variance in time t is now determined by an ARCH (p) model plus the sum of the conditional variances in the previous periods times their coefficients back to period t-q. Neither in this case are the parameters allowed to be negative. The stationary restriction is now that the sum of the \( \alpha_i \)'s
plus the sum of the $\beta_i$’s must be less than unity. The term $\sigma_t^2$ in the equations is in this thesis the conditional variance in the house prices deviation from the house price function. Bollerslev (1986) shows that a GARCH (1, 1) model equates with an infinite ARCH model. This implies that a GARCH model is better for capturing the persistence in the conditional variance than a low specified ARCH model. Many of the researchers mentioned in the introduction conclude that there is high persistence in the conditional variance of house prices. It is likely that this is also the case in this thesis. Thus a GARCH model is probably better for explaining the conditional variance than an ARCH model in this case. The GARCH model takes all previous shocks into account, while the ARCH model only takes the $p$ chosen lags into account.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2$$

(3)

The properties of an ARCH ($p$) model are similar to the ones for the GARCH ($p$, $q$). The steps when taking the unconditional variance for a GARCH ($p$, $q$) model are somewhat more complicated though. However, the important implication is that it is constant and independent of time. Equation (4) determines the residual in time $t$, $\varepsilon_t$. The residuals in the equations below are in this thesis those that come from the house price function of fundamental variables.

$$\varepsilon_t = v_t \sqrt{\sigma_t^2} = v_t \sigma_t$$

(4)

In Equation (4), $\sigma_t^2$ is the conditional variance and could be defined both as in Equation (2) and (3). The residuals are assumed to be white noise. It implies that no matter what $t$ is, they have a zero mean and constant variance, $\sigma$, and the covariance is zero between all residuals in different time periods. The variable $v_t$ is also assumed to be white noise with a variance equal to unity. All $v_t$’s and $\varepsilon_t$’s are assumed to be independent of each other. In Equation (5), the unconditional expected value is taken from Equation (4).

$$E(\varepsilon_t) = E(v_t \sigma_t) = E(v_t) * E(\sigma_t) = 0 * E(\sigma_t) = 0$$

(5)
Since the variables are independent of each other and the expected value of $v_t$ is zero, the expected value of $\varepsilon_t$ is zero. These steps and the outcome are exactly the same, no matter if $\sigma_t^2$ is an ARCH (p) or a GARCH (p, q) model. Taking the unconditional variance of Equation (4) is the same as taking the expectations of the squared $\varepsilon_t$. This because the covariance between all different $\varepsilon_t$’s is zero so all cross products between $\varepsilon_t$’s in different time periods will be zero. As mentioned above, the unconditional variance is constant and independent of time.

When taking the conditional expectation of Equation (4), the expectation of $\varepsilon_t$ is taken given its previous values. The history of $\varepsilon_t$ is defined as $\psi_{t-1}$. It stands for $\varepsilon_{t-1}$ and all its previous lags. The conditional expectation of Equation (4) is taken in Equation (6) below.

$$E(\varepsilon_t|\psi_{t-1}) = E(v_t|\psi_{t-1}) \ast E(\sigma_t|\psi_{t-1}) = E(v_t) \ast \sigma_t = 0 \quad (6)$$

Since $v_t$ and $\varepsilon_t$ are independent of each other, the conditional expectation of $v_t$ given $\psi_{t-1}$ is the unconditional expectation of $v_t$, $E(v_t)$. As shown in Equation (2), $\sigma_t$ is a function that depends on the lags of $\varepsilon_t$. Therefore the conditional expected value of $\sigma_t$ given $\psi_{t-1}$, is just the function itself, $\sigma_t$. Since $E(v_t)$ is equal to zero, the conditional expectation is also equal to zero. Hence the expected value is the same no matter if the unconditional or the conditional expectations are taken. The variance however, differs in the two cases. In Equation (7) below, the conditional variance is taken of Equation (4).

$$E(\varepsilon_t^2|\psi_{t-1}) = E(v_t^2|\psi_{t-1}) \ast E(\sigma_t^2|\psi_{t-1}) = E(v_t^2) \ast \sigma_t^2 = \sigma_t^2 \quad (7)$$

From the same reasoning as above, it can be concluded that the conditional variance of $\varepsilon_t$ given its entire history is equal to $\sigma_t^2$. The outcome is the same, no matter if $\sigma_t^2$ is an ARCH (p) or a GARCH (p, q) model. Thus it is shown that the conditional variance in the deviations of the house price function can be modeled by Equation (2) or (3).

4.2 TARCH and EGARCH models

Not only can there be ARCH and GARCH effects in the residuals of a function, there can also be asymmetric effects in those residuals. Two kinds of models are used to find out whether asymmetric effects exist. Those models are TARCH developed by Glosten et al. (1994) and
EGARCH developed by Nelson (1991). For simplicity the models are shown in their simplest GARCH form. The simplest GARCH model with a threshold term takes the form of Equation (8).

\[
\sigma_t^2 = \alpha_0 + \lambda_1 d_{t-1} \epsilon_{t-1}^2 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2
\]  

Equation (8) is a GARCH (1, 1) model plus the term \(\lambda_1 d_{t-1} \epsilon_{t-1}^2\), which is the threshold term of this TGARCH model. In the threshold term \(d_{t-1}\) is a dummy variable that takes the value of one if the shock is negative, otherwise it is zero. If the estimated coefficient, \(\lambda_1\), significantly differs from zero then one can conclude that asymmetric effects exist. If such effects do exist, Equation (8) should fit better with the data than an ordinary GARCH (1, 1) model would. If negative shocks occur they will have the effect of \((\lambda_1 + \alpha_1)\) times the squared lagged residual and positive shocks will only have the effect of \(\alpha_1\) times the squared lagged residual on the conditional variance. The researcher can choose to include more ARCH, GARCH or threshold terms in Equation (8).

The EGARCH model is a logarithmic GARCH model with a term that allows for asymmetric effects. Nelson (1991) argues that it is better to use lags of standardized residuals instead of normal squared ones. A standardized residual is the normal residual divided by the square root of the expression for the conditional variance, \(\sigma_t\). It is shown in Equation (9).

\[
s_t = \frac{\epsilon_t}{\sqrt{\sigma_t^2}} = \frac{\epsilon_t}{\sigma_t}
\]  

For simplicity also in this case, an EGARCH model in its simplest form is shown in Equation (10). The researcher can, as in the previous case, choose to include more terms. The simplest EGARCH model takes the form of Equation (10).

\[
\ln(\sigma_t^2) = \alpha_0 + \lambda_1 \frac{\epsilon_{t-1}}{\sigma_{t-1}} + \alpha_1 \frac{\epsilon_{t-1}}{\sigma_{t-1}} + \beta_1 \ln(\sigma_{t-1}^2)
\]  

Since this model is logarithmic on the left hand side, it does not have non-negative constraints for the coefficients like in the cases earlier described for standard ARCH or GARCH models. If there are any asymmetric effects those will be captured by \(\lambda_1\). If the estimated \(\lambda_1\)
significantly differs from zero the effect of a positive shock will be \((\lambda_1 + \alpha_1)\) and the effect of a negative shock will be \((\lambda_1 - \alpha_1)\). The most common asymmetric effect is the leverage effect, i.e. when negative shocks have greater impact on the conditional variance than positive ones of equal size. Like earlier mentioned, this is the case both when it comes to financial markets and house markets. Thus negative shocks often cause speculative herd behavior.

4.3 Testing for ARCH and GARCH effects and how to specify a model

Before the estimation of an ARCH or a GARCH model one should start by deciding if such a model is needed by testing for serial correlation in the squared residuals. There are two kinds of tests commonly used when testing for ARCH and GARCH effects. The first was worked out by McLeod and Li (1983). The first step in their test is to run OLS on the auxiliary regression in Equation (11) below.

\[ \varepsilon_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 \]  

The second step is to test:

\[
\left\{ \begin{array}{l}
H_0: \alpha_1 = \alpha_2 = \ldots = \alpha_q = 0 \\
H_1: \text{At least one of the } \alpha_i \text{'s differs from zero}
\end{array} \right.
\]

The test statistic is calculated by multiplying \(R^2\) times the number of observations. It converges to a chi-squared distribution with \(q\) degrees of freedom. For small samples usually the F-statistic is used with \(q/(T-q)\) degrees of freedom. If the null hypothesis is correctly accepted, it implies there is no serial correlation in the squared residuals. Thus no ARCH or GARCH effects exist. In that case the variance in the fluctuations from the house price function is constant and independent of time. Then house prices fluctuations from the function are of similar size on average in different time periods. This makes house prices predictable because the house price function is accurate. If the null hypothesis is correctly rejected, the opposite is true and thus ARCH or GARCH effects exist. It implies that unusually large shocks in the house price function cause increased conditional variance in the functions residuals. This makes house prices harder to explain by fundamental variables. In order to
better explain house prices, the residuals should be modeled by an ARCH or GARCH model. When the deviations from the function are greater than usual, the country is by definition in a house price bubble. Enders (2010) also suggests that one can use the Ljung-Box Q-test to find out if there is serial correlation in the squared residuals.

If the researcher concludes that ARCH or GARCH effects do exist, such a model should be specified. Enders (2010) argues that the researcher should begin with an ARCH (1) model and then test if there are any remaining effects the ARCH (1) model is not able to capture. If so, a higher order ARCH or GARCH model should be specified. When testing if there are any remaining ARCH or GARCH effects in the residuals one should start by obtaining the standardized and the squared standardized residuals. The next step is to test by Ljung-Box Q-test if there is serial correlation in the two series. As soon as the researcher estimates a model with no remaining ARCH or GARCH effects he or she should stop there.

When specifying an ARCH or GARCH model it is not only the ps and qs the researcher has to determine but also the assumed distribution of the residuals. These models are estimated by maximum likelihood and thus the assumption about the distribution of the residuals has a strong impact on the estimation process. Therefore a test of normality is preferred for guidance about the distribution. In this thesis the Jarque-Berra test is used for this purpose.

5. Results

5.1 Cointegration results

When it comes to macroeconomic time series variables these are often non-stationary, which is also the case for the variables included in Equation (1). To achieve a regression where the residuals are stationary, the variables must be cointegrated. If the variables are to be cointegrated they must be integrated in the same order. In Table 5.1 the results of Augmented Dickey-Fuller (ADF) tests for all variables are presented. They are presented both in their original form and when taking the first difference. From Figure 3.1 it is not clear in all cases what terms should be included in the ADF tests. Hence the ADF tests are performed both in the case with an intercept and time trend included, and in the case with only an intercept included. The results are presented in Table 5.1. The t-stat (1) is the statistic when only an intercept is included in the ADF test. The t-stat (2) is the statistic when both an
Table 5.1: ADF tests for all countries and variables in original and in first difference form.

**Panel (a): UK**

<table>
<thead>
<tr>
<th></th>
<th>$p_t$</th>
<th>$\Delta p_t$</th>
<th>GDP$_t$</th>
<th>$\Delta$GDP$_t$</th>
<th>$K_t$</th>
<th>$\Delta K_t$</th>
<th>pop$_t$</th>
<th>$\Delta$pop$_t$</th>
<th>$r_t$</th>
<th>$\Delta r_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-stat (1)</td>
<td>-1.286</td>
<td>-5.219</td>
<td>-1.133</td>
<td>-4.141</td>
<td>-0.589</td>
<td>-2.307</td>
<td>3.781</td>
<td>-1.041</td>
<td>0.155</td>
<td>-7.931</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.637)</td>
<td>(0.000)</td>
<td>(0.704)</td>
<td>(0.001)</td>
<td>(0.870)</td>
<td>(0.170)</td>
<td>(1.000)</td>
<td>(0.739)</td>
<td>(0.969)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>t-stat (2)</td>
<td>-1.455</td>
<td>-5.241</td>
<td>-1.155</td>
<td>-4.248</td>
<td>-2.397</td>
<td>-2.273</td>
<td>1.822</td>
<td>-3.448</td>
<td>-1.918</td>
<td>-7.959</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.843)</td>
<td>(0.000)</td>
<td>(0.917)</td>
<td>(0.004)</td>
<td>(0.381)</td>
<td>(0.447)</td>
<td>(1.000)</td>
<td>(0.047)</td>
<td>(0.643)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

**Panel (b): USA**

<table>
<thead>
<tr>
<th></th>
<th>$p_t$</th>
<th>$\Delta p_t$</th>
<th>GDP$_t$</th>
<th>$\Delta$GDP$_t$</th>
<th>$K_t$</th>
<th>$\Delta K_t$</th>
<th>pop$_t$</th>
<th>$\Delta$pop$_t$</th>
<th>$r_t$</th>
<th>$\Delta r_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-stat (1)</td>
<td>-2.811</td>
<td>-2.434</td>
<td>-0.103</td>
<td>-4.306</td>
<td>-1.272</td>
<td>-2.373</td>
<td>-0.590</td>
<td>-1.760</td>
<td>-1.232</td>
<td>-5.055</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.058)</td>
<td>(0.133)</td>
<td>(0.947)</td>
<td>(0.001)</td>
<td>(0.644)</td>
<td>(0.150)</td>
<td>(0.870)</td>
<td>(0.400)</td>
<td>(0.662)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>t-stat (2)</td>
<td>-3.853</td>
<td>-2.431</td>
<td>-2.067</td>
<td>-4.302</td>
<td>-2.738</td>
<td>-2.466</td>
<td>-2.412</td>
<td>-1.266</td>
<td>-2.101</td>
<td>-5.142</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.015)</td>
<td>(0.363)</td>
<td>(0.563)</td>
<td>(0.003)</td>
<td>(0.222)</td>
<td>(0.345)</td>
<td>(0.373)</td>
<td>(0.895)</td>
<td>(0.543)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

**Panel (c): Sweden**

<table>
<thead>
<tr>
<th></th>
<th>$p_t$</th>
<th>$\Delta p_t$</th>
<th>GDP$_t$</th>
<th>$\Delta$GDP$_t$</th>
<th>$K_t$</th>
<th>$\Delta K_t$</th>
<th>pop$_t$</th>
<th>$\Delta$pop$_t$</th>
<th>$r_t$</th>
<th>$\Delta r_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-stat (1)</td>
<td>0.179</td>
<td>-3.665</td>
<td>0.315</td>
<td>-4.736</td>
<td>-1.929</td>
<td>-1.627</td>
<td>1.123</td>
<td>-1.372</td>
<td>-0.844</td>
<td>-7.292</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.970)</td>
<td>(0.006)</td>
<td>(0.978)</td>
<td>(0.000)</td>
<td>(0.318)</td>
<td>(0.465)</td>
<td>(0.998)</td>
<td>(0.593)</td>
<td>(0.802)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.828)</td>
<td>(0.020)</td>
<td>(0.451)</td>
<td>(0.000)</td>
<td>(0.082)</td>
<td>(0.766)</td>
<td>(0.377)</td>
<td>(0.387)</td>
<td>(0.188)</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>
intercept and a time trend are included in the ADF test. The results for the UK are presented in Panel (a), for the USA in Panel (b) and for Sweden in Panel (c). Sweden’s tests are performed on the quarterly data because it stretches for a longer period than the monthly data.

For the UK, the test results are clear when it comes to house prices, GDP and the interest rate. In their original form, those series’ all contain a unit root, since the null hypothesis is accepted with high p-values. When taking the first difference, the null hypothesis is rejected and thus the ADF tests indicates that the variables are I(1). The results are the same no matter what terms are included in the ADF tests. For the population, the null hypothesis is clearly accepted in its original form. After the first difference is taken, it is rejected on the five percent level if both the time trend and intercept are included. The interpretation of those test results is that this series is also I(1). For the capital stock of housing, the null hypothesis is always accepted on conventional levels. However, Enders (2010) and Verbeek (2012) among others, argue that the power of the ADF test is low. It implies that the null hypothesis is accepted even though it should be rejected in many cases. When looking at Figure 3.1 again, it rather looks like there are structural breaks instead of the series not being I(1). Therefore it is assumed that the capital stock of housing is I(1) even though this is not in line with the results from the ADF tests. After all, most macroeconomic variables are I(1).

For the USA, the null hypotheses are accepted for the differentiated house prices, the capital stock of housing and the population. For the first two, it is likely to be for the same reasons as mentioned in the case of the UK. For the population it is hard to say. The series might not be I(1). If that is the case, this will probably not be problematic in the regression results anyway. The other American variables are clearly I(1) according to the ADF tests. In the Swedish case, all variables but the capital stock of housing and the population are I(1) according the ADF tests. In this case, the same reasoning applies as in the previous two cases. In summation, it is assumed that all variables for all countries to be I(1). There are some hesitations when it comes to the population series’, but if those are not I(1), they will not be very problematic for the estimation process. To avoid spurious regressions, the variables are tested for cointegration by the Engle-Granger test. The results are presented in Table 5.2. The variables given in the table are the dependent variables. The same panels apply for the same countries as in Table 5.2.
Table 5.2: Engle-Granger tests for cointegration for all countries

Panel (a): UK

<table>
<thead>
<tr>
<th>tau-statistic</th>
<th>p_t</th>
<th>GDP_t</th>
<th>K_t</th>
<th>pop_t</th>
<th>r_t</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p-value)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1.542</td>
<td>-2.157</td>
<td>-0.999</td>
<td>-1.288</td>
<td>-3.993</td>
<td></td>
</tr>
<tr>
<td>(0.983)</td>
<td>(0.914)</td>
<td>(0.996)</td>
<td>(0.992)</td>
<td>(0.143)</td>
<td></td>
</tr>
</tbody>
</table>

Panel (b): USA

<table>
<thead>
<tr>
<th>tau-statistic</th>
<th>p_t</th>
<th>GDP_t</th>
<th>K_t</th>
<th>pop_t</th>
<th>r_t</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p-value)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-2.545</td>
<td>-3.319</td>
<td>-4.457</td>
<td>-2.736</td>
<td>-2.172</td>
<td></td>
</tr>
<tr>
<td>(0.801)</td>
<td>(0.421)</td>
<td>(0.049)</td>
<td>(0.721)</td>
<td>(0.911)</td>
<td></td>
</tr>
</tbody>
</table>

Panel (c): Sweden

<table>
<thead>
<tr>
<th>tau-statistic</th>
<th>p_t</th>
<th>GDP_t</th>
<th>K_t</th>
<th>pop_t</th>
<th>r_t</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p-value)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-3.940</td>
<td>-5.086</td>
<td>-4.433</td>
<td>-4.069</td>
<td>-5.097</td>
<td></td>
</tr>
<tr>
<td>(0.176)</td>
<td>(0.013)</td>
<td>(0.066)</td>
<td>0.140</td>
<td>(0.013)</td>
<td></td>
</tr>
</tbody>
</table>
For the UK the tests indicate no signs of cointegration. In the American case cointegration is barely found on the five percent level with the capital stock of housing as a dependent variable. When using any other dependent variable, the tests clearly fail to find cointegration. For Sweden, cointegration is found on the five percent level when using either GDP or the interest rate as a dependent variable. When letting the capital stock of housing be the dependent variable, cointegration is found on the ten percent level. In all other cases, cointegration cannot be found even though it is much closer than for the other two countries. When letting the house price be the dependent variable, cointegration is never detected on conventional levels for any country. This implies that in order for the residuals in Equation (1) to be stationary, all variables must first be differentiated one time. Therefore Equation (12) is used as mean function in this thesis.

\[ \Delta \ln(p_t) = \beta_0 + \beta_1 \Delta \ln(GDP_t) + \beta_2 \Delta \ln(K_t) + \beta_3 \Delta \ln(pop_t) + \beta_4 \Delta r_t + \varepsilon_t \quad (12) \]

Equation (12) is, like Equation (1), a benchmark where lags can be added if it improves the fit of the model. In Equation (12) it is only the interaction between the variables on short sight that is described. One must have this in mind when interpreting the estimates of this equation. All differentiated variables are presented in the appendix.

5.2 Regression results

The results from the regressions of the house price functions are shown in Table 5.3. One regression is run for the UK, one for the USA, one for Sweden based on the quarterly data (denoted as Sweden (Q)) and one for Sweden based on the monthly data (denoted as Sweden (M)). All variables are included in the regressions even though some of them do not significantly differ from zero on conventional levels. In addition, lags of some variables are included since they improve the fit of the model with the data. It implies that the two chosen information criteria, AIC and SBC, take on lower values. For the UK and Sweden (Q), a lag of the house price is included. For the USA, two lags of the house price and one lag of the capital stock of housing are included. Additionally a dummy variable is included in that regression. The dummy variable corrects for a structural break that occurs in about year 1998. The dummy thus take the value of one from year 1998 and onwards and is otherwise equal to zero. For the USA structural breaks also occur during the latest financial crisis. It was not
possible to correct for those and therefore the time series is shortened so that only stretches to 2006. For the UK it can be concluded from a CUSUM square test that a structural break occurs in the beginning of the sample. Unfortunately it was not possible to correct for it in any way. All CUSUM tests for the regressions run in Table 5.3 are presented in the appendix.

In Table 5.3 there are two insignificant parameters for the UK, the change of the population and the interest rate. Both of these variables’ coefficients also exhibit the opposite sign than expected. If the population grows it is reasonable to assume, other things equal, that house prices will go up. Since the estimated coefficient for population growth is negative not only for the UK, but in all other regressions as well, this is obviously not the case. Perhaps houses are built at a faster pace than population growth so the demand for houses actually goes down while the population grows. The sign of the coefficient for the interest rate is also opposite from expectations. The same is true for all cases but for Sweden (M). One possible explanation is that there is a lag in the effect on house prices from a change in interest rates. Then it does not affect house prices in the same period and thus the estimated coefficient is of the wrong sign and insignificant. When lags of the interest rate are included their coefficients are still insignificant. Another possible explanation is that interest rates do not affect house prices in the short run. This seems reasonable since it is highly insignificant in almost all cases.

All other coefficients differ significantly from zero on the one percent level for the UK. The sign of GDP’s coefficient is as expected. When it comes to the coefficient for the capital stock of housing there were two possible expectations. It makes sense to expect a negative sign, because if the supply of housing goes up, house prices should go down, other things equal. On the other hand it is also reasonable to assume a positive sign, especially when looking at Figure 3.1. There one can observe that the growth of the house stock behaves like a smooth trend for house prices in the case of UK. If house prices go up, there is a greater possibility for building companies to sell houses for high prices and thus bigger possibilities for high profits. This could encourage the companies build more and hence a positive sign should be expected. Since the sign of the coefficient is positive in all cases, the latter reasoning seems to be the most likely to be the correct one. In the regression for UK, 31.4% of the variation in house prices is explained by this regression in the short run.
For the USA, 62.7% of the variation of house prices is explained by the regression. The intercept is not significant this time and the coefficient for GDP has a negative sign, which is not in line with the expectations. One possible explanation for this is due to the fact that the house price is a leading indicator of GDP. Another difference from the case of the UK is that there are three additional variables. All three variables’ coefficients are significant on the one percent level.

Table 5.3: Results for the mean equations

<table>
<thead>
<tr>
<th></th>
<th>UK</th>
<th>USA</th>
<th>Sweden (Q)</th>
<th>Sweden (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>β₀</strong></td>
<td>-0.006***</td>
<td>0.001</td>
<td>0.001**</td>
<td>0.006**</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.005)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Δln(GDPᵣ)</td>
<td>1.000***</td>
<td>-0.121**</td>
<td>0.351*</td>
<td>0.480*</td>
</tr>
<tr>
<td></td>
<td>(0.270)</td>
<td>(0.059)</td>
<td>(0.191)</td>
<td>(0.267)</td>
</tr>
<tr>
<td>Δln(Kᵣ)</td>
<td>2.672***</td>
<td>6.670***</td>
<td>0.252</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(0.455)</td>
<td>(1.178)</td>
<td>(0.310)</td>
<td>(0.914)</td>
</tr>
<tr>
<td>Δln(Kᵥ₋₁)</td>
<td></td>
<td>-5.542***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.117)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δln(popᵣ)</td>
<td>-0.505</td>
<td>-2.575**</td>
<td>-6.410**</td>
<td>-6.587</td>
</tr>
<tr>
<td></td>
<td>(3.997)</td>
<td>(1.065)</td>
<td>(3.018)</td>
<td>(6.441)</td>
</tr>
<tr>
<td>Δrᵣ</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>-0.003*</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Δln(Pᵥ₋₁)</td>
<td>0.210***</td>
<td>0.518***</td>
<td>0.584***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.053)</td>
<td>(0.087)</td>
<td></td>
</tr>
<tr>
<td>Δln(Pᵥ₋₂)</td>
<td></td>
<td>-0.150***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.053)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dᵣ</td>
<td></td>
<td>0.001***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.314</td>
<td>0.627</td>
<td>0.527</td>
<td>0.046</td>
</tr>
<tr>
<td>F-statistic</td>
<td>32.347***</td>
<td>77.512***</td>
<td>23.490***</td>
<td>2.773**</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.006</td>
<td>1.994</td>
<td>2.226</td>
<td>2.054</td>
</tr>
<tr>
<td>AIC</td>
<td>-6.389</td>
<td>-9.223</td>
<td>-5.604</td>
<td>-6.124</td>
</tr>
<tr>
<td>SBC</td>
<td>-6.322</td>
<td>-9.126</td>
<td>-5.449</td>
<td>-6.023</td>
</tr>
</tbody>
</table>

In the case of Sweden (Q), the same variables are included as in the case for UK. All coefficients have the same signs. The coefficient for GDP is now only significant on the 10% level and the coefficient for the capital stock of housing is not significant on conventional levels. In the data section, low growth in the housing stock was noticed for Sweden, so it is not a surprising result. As in the US regression, the coefficient for population is also significant on a five percent level. Sweden’s (Q) regression explains 52.7% of the variation in
house prices. For Sweden (M) only 4.6% of variation in the house prices is explained by the regression. The signs of the coefficients are the same for all variables as in the case of Sweden (Q) but for the interest rate. This time the interest rate is actually significant on the 10% level. Another difference from the previous Swedish case is that the coefficient for the population growth now becomes insignificant. The latest regression’s inability to explain the house prices is startling and could be a sign that Sweden has entered a house price bubble or not. However, it is beyond the scope of this thesis to determine whether a country is in a house price. It could also be the case that the series’ are simply too short.

To answer the question, if there are conditional heteroskedastic effects for house prices deviations from the house price function, tests for those effects are made. The results of the tests are presented in Table 5.4 below. Two tests are performed, the Ljung-Box Q-test and the LM-test of the squared residuals. The results are presented with four and eight lagged residuals, just like Tse (1998) does with standardized residuals. The Qs stand for a Q-test, the subscript two symbolizes the test is carried out on squared residuals. The number in the parenthesis displays the number of lags included in the test. For the LM-test, the F-statistics and their p-values are presented. As mentioned in the method section, F-statistics are only used for small samples. Unfortunately, the econometric software E-views only provides the p-value for the χ²-statistic and not the χ²-statistic itself. Therefore only the results from the F-test are shown. Since the p-values from the F-tests and the χ²-tests are almost identical, this does not affect the results.

<table>
<thead>
<tr>
<th></th>
<th>UK</th>
<th>USA</th>
<th>Sweden (Q)</th>
<th>Sweden (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p-value)</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.161</td>
</tr>
<tr>
<td>(p-value)</td>
<td>0.002</td>
<td>0.000</td>
<td>0.006</td>
<td>0.338</td>
</tr>
<tr>
<td>F-stat (4)</td>
<td>4.075</td>
<td>4.657</td>
<td>3.394</td>
<td>1.606</td>
</tr>
<tr>
<td>(p-value)</td>
<td>0.003</td>
<td>0.001</td>
<td>0.012</td>
<td>0.176</td>
</tr>
<tr>
<td>F-stat (8)</td>
<td>2.265</td>
<td>3.180</td>
<td>1.800</td>
<td>0.864</td>
</tr>
<tr>
<td>(p-value)</td>
<td>0.023</td>
<td>0.002</td>
<td>0.088</td>
<td>0.549</td>
</tr>
</tbody>
</table>

For UK, the results are clear. In all four cases the test rejects the null hypothesis and thus it can be concluded that there are conditional heteroskedastic effects in the residuals. For USA the results are even clearer. In all four cases the null hypothesis can be rejected and hence the
same conclusion can be drawn for USA. For Sweden (Q), both Q-tests clearly reject the null hypothesis. With four lags the LM-test clearly rejects the null hypothesis too, but the p-value on 8.8% for the test with eight lags is not equally convincing. Even though the rejection is not clear-cut in the latter case, the interpretation of the results is that conditional heteroskedastic effects exist in the residuals for Sweden (Q). One must not forget that the data is given on a quarterly basis and hence those effects are not as easily detected as in the case with data given on a monthly basis. If the data for this period would have been on a monthly basis, the tests would probably have shown clearer results. For Sweden (M) there is no rejection of the null hypothesis on conventional significance levels in any test. This result was not unexpected since the regression did not manage to explain much of the variation in the house prices.

Before the estimation of the ARCH and GARCH models, Jarque-Bera tests are performed. The reason is to test if the residuals of the above regressions are normally distributed. The test is not made for Sweden (M) because no ARCH effects were found. ARCH and GARCH models are estimated by Maximum Likelihood and thus the assumption of the residuals distribution is of importance. In the Jarque-Bera test, the following hypothesis is tested,

\[
\begin{align*}
H_0 &: \text{The residuals are normally distributed} \\
H_1 &: \text{The residuals are not normally distributed}
\end{align*}
\]

The results from the tests are presented in Table 5.5 below

<table>
<thead>
<tr>
<th></th>
<th>UK</th>
<th>USA</th>
<th>Sweden (Q)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JB-statistic</td>
<td>53.999</td>
<td>19.002</td>
<td>5.549</td>
</tr>
<tr>
<td>(p-value)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.062)</td>
</tr>
</tbody>
</table>

In the first two cases, for the UK and USA, there is a clear rejection of the null hypothesis. For Sweden (Q) the null hypothesis is accepted on a significance level of five percent. Thus it is better to assume t-distributed residuals for the UK and USA and normally distributed ones for Sweden (Q).

The estimations of the ARCH and GARCH models are presented in Table 5.6. In the table, estimations are presented with two distribution assumptions of the most appropriate ARCH or GARCH model for each country. In the decision of which assumption best fits with the data,
the two information criteria and the log Likelihood are compared between the models. Ljung-Box Q-tests for the standardized residuals are also presented in the table in order to decide if there are any remaining ARCH or GARCH effects. The subscript one displays that the test is on the non-squared standardized residuals. The subscript two signifies the test is on the squared standardized residuals. As in previous cases, the number in the parenthesis denotes the number of lags included in the test.

In Table 5.6 the best ARCH and GARCH models are presented for each country. The models with assumed normal distribution of the residuals are estimated with Bollerslev-Wooldridge robust standard errors and covariance. This is standard practice in these kind of models. For example Miller & Peng (2006) and Morley & Thomas (2010) do the same thing. In the decision of choosing these particular models, I look if there are any remaining ARCH or GARCH effects in the standardized and the squared standardized residuals. As soon as those effects are not detected by the Ljung-Box Q-tests (i.e. accepting the null hypothesis), that model is chosen. Then the model fulfills its aim, to explain the ARCH or GARCH effects, and is as parsimonious as it gets. Unfortunately in the case of the UK, a model that accepts the null hypothesis for the standardized residuals is never found. Nevertheless the best fitting model for the UK is shown in the table. For the USA, the null hypothesis is always accepted. In the case of Sweden, the null hypothesis is always accepted on a significance level of five percent. The conclusions drawn from the American and Swedish results are that the model captures the ARCH and GARCH effects.

The different values of logL, AIC and SBC between the two models for each country, are in line with the conclusions drawn from the Jarque-Berra tests. For the UK and USA all three variables improve by changing the assumption from normal distributed residuals to t-distributed ones. For Sweden logL and AIC is approximately the same in the two cases but SBC improves significantly by changing the assumption from t-distribution to normal distribution. Hence it can be concluded that the best model for the UK and USA are those with the assumption of t-distributed residuals. For Sweden it is the model with residuals assumed to be normally distributed.

The model for the UK has two significant parameters, $\alpha_1$ on the ten percent level and $\beta_1$ on the one percent level. One should note that the sum of these estimated coefficients is greater than unity, which implies non-stationarity. The effects of a shock in this model are not only
permanent, but also explosive. In the case of the USA, $\alpha_1$ is significant on a five percent level and $\beta_1$ is significant on a one percent level.

They sum up to 0.9735 which is very close to unity. In similarity Miller & Peng (2006) get a sum of those coefficients of about 0.98. They argue that it is still stationary but the adjustment process is slow. This implies that if a shock occurs, there follows quite long period of higher conditional variance. In the Swedish case, the intercept term in the ARCH (1) model, $\alpha_0$, is significant on a one percent level even though the estimate is close to zero. The ARCH term, $\alpha_1$, is significant on a ten percent level. The estimate of $\alpha_1$, on 0.256, does not imply such persistence as in the previous two cases. You never get the same persistence in an ARCH model as in a GARCH model. However, the conditional variance in the Swedish model might be high during times of turbulence anyway. When financial crises take place, large shocks usually occur and these tend to create a series of following big shocks. In the American model, the high persistence should be a warning for policy makers. The model indicates that in times of crises, the conditional variance increases. It not only increases, but also remains on higher levels than usual. The greater amplitude of the shocks and the longer their effects last, the greater damage they cause the economy.

Table 5.6: ARCH and GARCH models of the residuals

<table>
<thead>
<tr>
<th></th>
<th>UK GARCH (1, 1)</th>
<th>USA GARCH (1, 1)</th>
<th>Sweden ARCH (1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>t-dist</td>
<td>Normal</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-0.000</td>
<td>-0.001</td>
<td>0.000</td>
</tr>
<tr>
<td>$\alpha_0$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.071**</td>
<td>0.099*</td>
<td>0.089**</td>
</tr>
<tr>
<td></td>
<td>0.033</td>
<td>0.052</td>
<td>0.041</td>
</tr>
<tr>
<td>$\log L$</td>
<td>1112.847</td>
<td>1120.865</td>
<td>1704.756</td>
</tr>
<tr>
<td>$Q_1$(4)</td>
<td>16.701***</td>
<td>16.469***</td>
<td>1.016</td>
</tr>
<tr>
<td>$Q_2$(8)</td>
<td>32.076***</td>
<td>31.666***</td>
<td>4.245</td>
</tr>
<tr>
<td>$Q_2$(4)</td>
<td>1.264</td>
<td>1.546</td>
<td>1.083</td>
</tr>
<tr>
<td>$Q_2$(8)</td>
<td>3.009</td>
<td>3.296</td>
<td>3.889</td>
</tr>
</tbody>
</table>
Then what about the asymmetric effects of the shocks? In the estimation of TARCH and EGARCH models asymmetric effects are not found. The coefficient $\lambda_1$ is never significant on conventional levels. However this does not necessarily mean that these kinds of effects do not exist. Especially not since both Karaglou et al. (2011) and Morley & Thomas (2010) find them in American and British house price data. In the American case in this thesis, the latest financial crisis with the greatest fall in house prices during the sample is excluded. Thus the most likely period to detect asymmetric effects is also excluded.

The answer to the main question in this thesis is: yes, there are conditional heteroskedastic effects in the deviation from a fundamental function that explains house prices. At least in the short run. This thesis therefore contributes to debate on house prices. Those effects should be included not only in the answer of the question if a country is in a house price bubble or not. They should also be included in the analysis of the possible severity if a house price bubble bursts and what actions should be taken to prevent and solve such problems. This should be of importance for policy makers, central banks, and commercial banks. It is also important for present and future house owners to be aware of the shifting riskiness of owning a house.

The results and conclusions in this thesis are in line with many of those provided by other researchers. Tsai et al. (2010) have a similar mean function as in this thesis. Their mean function is an AR(1) process of the differentiated house prices in the UK. In addition I include the fundamental variables with extra explanatory power. All other cited papers also use some form of ARMA model of differentiated house prices as mean function. The addition of the fundamental variables explains house prices better than just an ARMA model. This improves both $R^2$ and the two information criteria compared to the case of just an AR model. Thus it is those residuals that should be analyzed. Furthermore Tsai et al. (2010) also estimate an incorrectly specified GARCH (1, 1) model where the sum of $\alpha_1$ and $\beta_1$ exceeds unity. According to the authors, this is probably due structural breaks that give regime shifts in the conditional variance. To be able to estimate a stable model, they find three different states of conditional volatility and estimate a SWARCH model with better outcome.

Additionally Karaglou et al (2011), Morley & Thomas (2009) and Miller & Peng (2006) point out the problem with structural breaks in the British and American data. These structural breaks are likely the biggest source of error in the British model in this thesis. They are also the reason why the latest five years cannot be included in the American sample. Another
suggestion for modeling the high persistence of house prices is to estimate a so called component GARCH model (CGARCH). This suggestion comes from Miles (2009). He argues that a standard ARCH or GARCH model is not able to capture the high persistence well and thus they should not be used for trying to do so.

The policy implications here suggested are also in line with those of other researchers. Since conditional heteroskedasticity does exist, the riskiness of owning a house might be a very risky investment at a certain point of time. House investments usually stretch for a long time. As a consequence it is very likely for a house owner to be exposed to increased volatility in house prices and thus the increased risk of owning a house. With this reasoning, home owners should increase their awareness of the riskiness of their investment. People who are about to buy a house should have this in mind. One cannot live on the margin, because if interest rates suddenly go up people will start to default on their loans. House prices then fall, many of the owners then have to sell their houses for a lower price and large reductions in wealth become fact, both on the micro and the macro level.

The commercial banks and other mortgage lenders should also be aware of this risk. Karaglou et al. (2011), Morley & Thomas (2010) and Miller & Peng (2006) points out the importance of commercial banks seeing lending for house investments more risky than they currently do. Morley & Thomas (2010) argue that the housing market in general has important effects on the macroeconomy, especially when it comes to wealth effects and the mortgage markets. In the paper by Miller & Peng (2006), they conclude that house price volatility Granger-cause other variables in the economy. With this in mind, it is also important to point out the importance of central banks and other policy makers recognizing the riskiness of the housing markets. They must consider the conditional variance when analyzing the house market situation. When unusually large shocks occur, the conditional variance of the house price function increases. Greater deviations from the house price function imply that house prices are now worse explained by fundamental variables. One likely explanation is that house prices are more driven by expectations than before. The increased volatility and speculative behavior is in some cases persistent and often harms the economy.

More research is needed in this area, which almost all of the authors point out in the cited papers about house prices and their conditional variance. I recommend other researchers to include fundamental variables in the mean function as these should not be excluded. If they
are differentiated, as in my case, it will probably raise both $R^2$ and the fit with the data. As compared to the case when only an ARMA model is used. It would also be interesting to analyze the conditional variance of the residuals from a function with fundamental variables that are not differentiated. Then cointegration between the variables must be found in order for them to be stationary. Strong evidence for cointegration between my chosen variables is not found, but cointegration between fundamental variables and house prices has been found by other researchers. Clausen et al. (2011) being one example. To deal with the problems of structural breaks and long persistence in the conditional variance for house prices, some other models than standard ARCH or GARCH model might be preferred. It could be a SWARCH model suggested by Tsai et al. (2010) or a CGARCH model suggested by Miles (2009).

6. Concluding remarks

The answer of the main question in this thesis is yes, conditional heteroskedastic effects do exist in the residuals of a function of fundamental variables that explain house prices. However, it is not the yes anticipated when work on this thesis was begun. The thought was to look at the residuals from a cointegrated relationship, i.e. the residuals from a function that holds in the long run. Strong evidence for cointegration between the variables was not found, which forced me to differentiate them in order to get stationary residuals. A function with differentiated variables means that only the interaction on short sight is analyzed. Of course, this is also interesting and this kind of function explains house prices better than ARMA functions used by other researchers. The fact that the variables are differentiated makes it easier to compare the results in this thesis with those provided by other researchers. In comparison, the inclusion of fundamental variables improves both the explanatory power and the fit of the models.

Not much has been done in the research of house prices and their conditional variance. More research is definitely needed. This fact is pointed out by the authors of almost every paper in this area. The conditional variance of the residuals from a house price function with fundamental variables has not been analyzed before. To use such a function as a mean function should be more appropriate than just an ARMA function of the house prices. It is the conditional variance in the deviations from a house price function that needs to be explained not that of the house price itself. If the conditional variance in only the house prices is
analyzed, one does not know if the change is due to fundamental variables or speculative behavior. If the fundamental variables are added in the analysis, one can analyze the changing conditional variance that is not driven by those. To deal with the problems of structural breaks, it is probably more suitable to use a model like SWARCH or CGARCH rather than a standard ARCH or GARCH model.

This research area is important since it has been shown that the volatility of house prices affects the performance of other macroeconomic variables. Especially mortgage markets and wealth are affected. It is important that commercial banks and present and future house owners see the true riskiness of a house investment. It could also be important in the process of deciding whether a country is in a house price bubble or not. Most importantly, increased volatility often harms the economy and is sometimes highly persistent. Thus it must be modeled appropriately in order for policy makers to take the right decisions in preventing and solving the problems of unstable house markets. Therefore it should lie in the interest of central banks to pay attention to this research area.
7. References


**Data Sources**

Halifax

OECD

Statistics Sweden

Freddie Mac

World Bank Database
8. Appendix

Figure 7.1: First difference of all logged variables for UK (interest rates are not logged)

Panel (a): House price

Panel (b): GDP

Panel (c): Capital stock of housing

Panel (d): Population

Panel (e): Interest rate
Figure 7.2: First difference of all logged variables for USA (interest rates are not logged)

Panel (a): House price  
Panel (b): GDP  
Panel (c): Capital stock of housing

Panel (d): Population  
Panel (e): Interest rate
Figure 7.3: First difference of all logged variables for Sweden (interest rates are not logged)

Panel (a): House price

Panel (b): GDP

Panel (c): Capital stock of housing

Panel (d): Population

Panel (e): Interest rate
Figure 7.4: CUSUM tests for the house price functions

Panel (a): UK

Panel (b): USA without dummy

Panel (c): USA with dummy variable

Panel (d): Sweden (Q)

Panel (e): Sweden (M)
Figure 7.5: CUSUM square tests for the house price functions

Panel (a): UK
Panel (b): USA without dummy variable
Panel (c) USA with dummy variable

Panel (d): Sweden (Q)
Panel (e) Sweden (M)